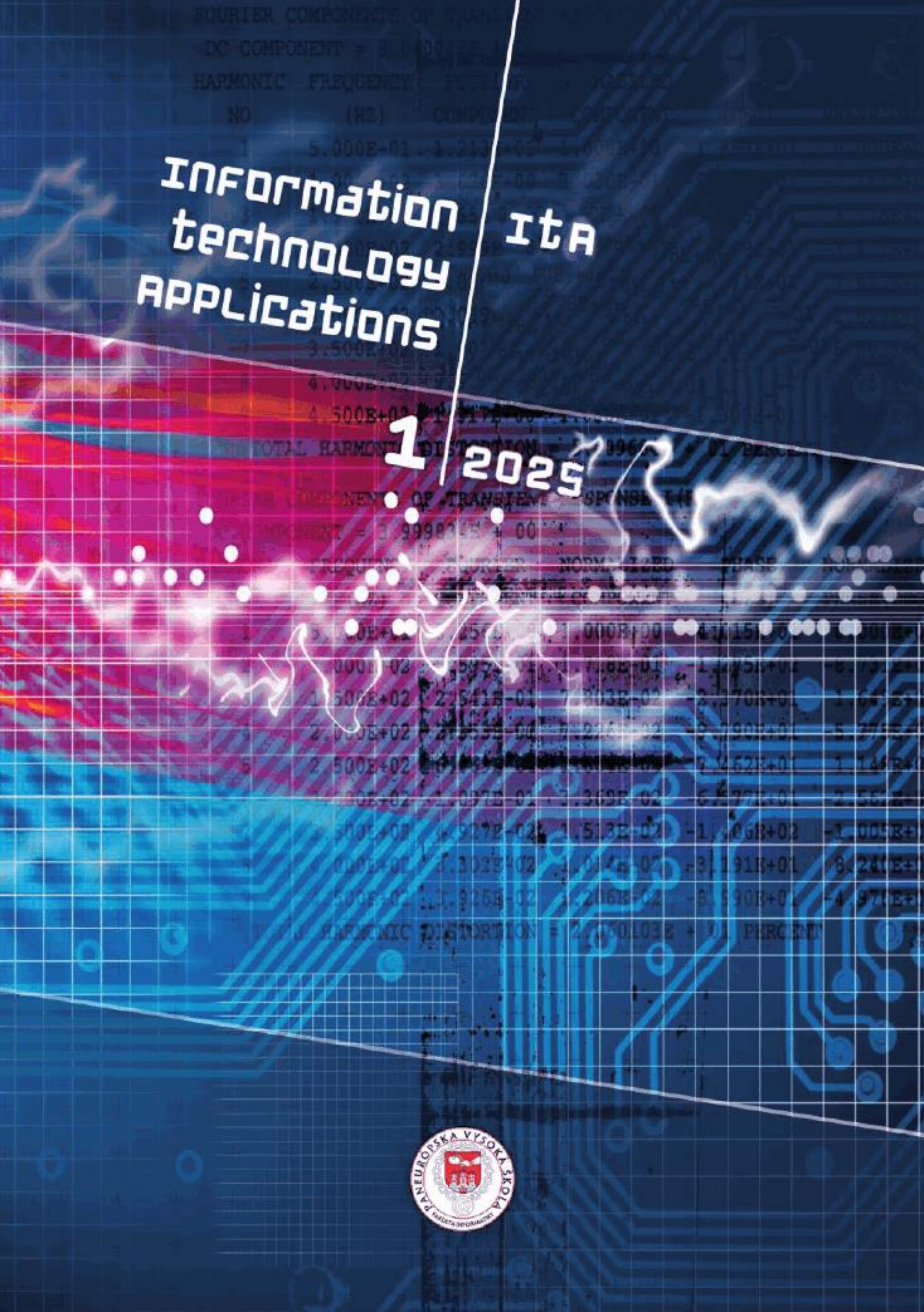
Information technology Applications

ItA

1 2025



PANEURÓPSKA VYSOKÁ ŠKOLA

PAN-EUROPEAN UNIVERSITY
Faculty of Informatics

VZDELÁVANIE VEDA VÝSKUM
EDUCATION SCIENCE RESEARCH

SSKI
SLOVENSKÁ SPOLOČNOSŤ PRE KYBERNETIKU A INFORMATIKU PRI SAV
SLOVAK SOCIETY FOR CYBERNETICS AND INFORMATICS

ITA - International Journal of Information Technology Applications

Volume 14, Number 1, July 2025

AIMS AND SCOPE OF ITA

The primary aim of the International Journal of Information Technology Applications (ITA) is to publish high-quality papers of new development and trends, novel techniques, approaches and innovative methodologies of information technology applications in the broad areas. The International Journal of ITA is published twice a year. Each paper is refereed by two international reviewers. Accepted papers will be available online with no publication fee for authors. The International Journal of ITA is being prepared for the bibliographic scientific databases.

Editor-in-Chief

prof. Ing. Štefan Kozák, PhD. Faculty of Informatics, Pan-European University in Bratislava
stefan.kozak@paneurouni.com

Executive Editor

Ing. Juraj Štefanovič, PhD., Faculty of Informatics, Pan-European University in Bratislava
juraj.stefanovic@paneurouni.com

Editorial Board

Mikhail A. Basarab, Russia
Ivan Brezina, Slovakia
Yakhua G. Buchaev, Russia
Ivana Budinská, Slovakia
Ján Cigánek, Slovakia
Ladislav Cvetko, Croatia
Gabriela Czanner,
Silvester Czanner,
United Kingdom
Zbigniew Domański, Poland
Maria Jose Escalona, Spain
Andrej Ferko, Slovakia
Vladimír S. Galayev, Russia
Pedro Gamito, Portugal
Ladislav Hudec, Slovakia
Oto Haffner, Slovakia
Matthias Harders, Austria
Adam Herout, Czech Republic
Ladislav Hluchý, Slovakia
Oleg Choporov, Russia
Aleš Janota, Slovakia
Mike Joy, United Kingdom
Gabriel Juhás, Slovakia
Martin Juhás, Slovakia
Jozef Kelemen, Czech Republic
Sergey Kirsanov, Russia
Vladimir I. Kolesnikov, Russia
Štefan Kozák, Slovakia
Alena Kozáková, Slovakia
Vladimír Krajčík, Czech Republic
Erik Kučera, Slovakia
Jaroslav Kultán, Slovakia
Ján Lacko, Slovakia
Igor Lvovich, Russia
Eva Mihaliková, Slovakia
Branislav Mišota, Slovakia
Pavle Mogin, New Zealand
Mykola S. Nikitchenko, Ukraine
Ján Paralič, Slovakia
Rimantas Petrauskas, Lithuania
Lucia Pomello, Italy
Martin Potančok, Czech Republic
Roman Povalej, Germany
Wolf Rauch, Austria
Danica Rosinová, Slovakia
Eugen Ružický, Slovakia
Václav Řepa, Czech Republic
Peter Sinčák, Slovakia
Elena Somova, Bulgaria
Vjeran Strahonja, Croatia
Jiří Šafařík, Czech Republic
Petr Šaloun, Czech Republic
József K. Tar, Hungary
Valentino Vranić, Slovakia
Adam Wojciechowski, Poland
Doina Zmaranda, Romania

ITA - International Journal of Information Technology Applications

Instructions for authors

The International Journal of Information Technology Applications is welcoming contributions related with the journal´s scope. Scientific articles in the range approximately 10 standard pages are reviewed by two international reviewers. Reports up to 5 standard pages and information notices in range approximately 1 standard page are accepted after the decision of editorial board. Contributions should be submitted via e-mail to the editorial office. The language of contributions is English. Text design should preserve the layout of the template file, which may be downloaded from the webpage of journal. Papers in this journal are provided under diamond acess policy (no fee, open website) and authors declare their own Creative Commons license.

Deadlines of two standard issues per year (special issues are possible)

paper submission deadline	– continuous process
review decision	– continuous process
release date	– July/December

Editorial office address

Faculty of Informatics, Pan-European University, Tematínska 10, 851 05 Bratislava, Slovakia
juraj.stefanovic@paneurouni.com

Published by

Pan-European University, Slovakia, https://www.paneurouni.com/
Paneurópska vysoká škola, n.o., Tomášikova 20, 821 02 Bratislava, IČO 36 077 429

Civil Association EDUCATION-SCIENCE-RESEARCH, Slovakia, https://v-v-v.eu/
OZ VZDELÁVANIE -VEDA-VÝSKUM, Andrusovova 5, 851 01 Bratislava,
IČO 42 255 180

Slovak Society for Cybernetics and Informatics (SSKI) at the Slovak Academy of Sciences, Slovakia
Slovenská spoločnosť pre kybernetiku a informatiku pri SAV (SSKI),
Ústav automobilovej mechatroniky, Fakulta elektrotechniky a informatiky STU,
Ilkovičova 3, 812 19 Bratislava 1, IČO 00 178 730, https://www.sski.sk/

Electronic online version of journal

https://www.itajournal.com/
(Open Journal System)

visit Archive and Instructions for authors:

Electronic backup and preservation

www.webdepozit.sk

digitálne pramene

Print version

Multigrafika s.r.o., Rajecká 13, 821 07 Bratislava

Print issues

Contact the editorial office,
print issues are available until they are in stock.

ISSN: 2453-7497 (online)
ISSN: 1338-6468 (print version)

Registration No.: EV 4528/12

Contents

Editorial

Dear readers,

The last article in this issue contains a short report about the 32^{th} International Conference Cybernetics & Informatics 2025, organised by Slovak Society of Cybernetics and Informatics in Mikulov, Czechia. The conference focuses on latest development in control and information technologies, related to multidisciplinary fields and industry.

First article is an another summary of running research specialized to early detection of neurodegenerative diseases by speech recording and analysis. Next four articles are reports from successful research works of our students at the end of their study. The last article reports about the construction of industrial system for control and data collection.

We will thank you for contributions - we welcome your latest research results and technical solutions and we look forward to further collaboration.

Ing. Juraj Štefanovič, PhD.
ITA Executive editor

International Journal
Information Technology Applications (ITA)
Volume 14, Number 1, June 2025

DETECTION OF NEURODEGENERATIVE DISEASES FROM SPEECH USING MULTIPLE CLASSES

Barbora Jurkovcová, Eugen Ružický, Štefan Kozák, Ján Lacko, Alfréd Zimmermann

Abstract:

Early detection of neurodegenerative diseases such as Alzheimer's disease (AD), mild cognitive impairment (MCI), and Parkinson's disease (PD) is crucial for timely intervention and improved patient outcomes. This study presents a machine learning framework for non-invasive diagnosis using voice signals recorded via smartphones. A machine learning model was developed, trained on the Slovak EWA-DB database, and evaluated on a large cohort consisting of patients with AD-MCI, PD, and healthy controls (HC). This model uses multi-class classification, which is more challenging than binary classification. The model achieved results with 93.5% accuracy, 93.6% sensitivity, and an F1 score of 88.5%, which are comparable to the results obtained from the EWA-DB for binary classification tasks. These results were achieved through thorough data preprocessing, including stratified sampling by age and diagnosis, class balancing through synthetic oversampling of minority classes, and dimensionality reduction through principal component analysis while preserving key information. We plan to apply the results independently of the multi-category classification of AD, MCI, and HC for non-invasive screening strategies in clinical practice. The proposed approach highlights the potential of speech biomarkers in combination with machine learning to improve early diagnosis across multiple classes.

Keywords:

Artificial intelligence, machine learning, smartphone application, EWA-DB database, detection of neurodegenerative diseases.

Introduction

Digital biomarkers represent a revolutionary tool in clinical decision-making, as they enable accurate, objective and continuous measurement of patients' health status using smartphones and wearable devices. These technologies can detect subtle motor, cognitive and speech changes that are often imperceptible by traditional methods. Speech signals such as changes in rhythm, pitch or articulation are showing promise in the diagnosis of neurological disorders such as Alzheimer's disease (AD) and Parkinson's disease (PD). Thanks to artificial intelligence and machine learning, these features can be analysed automatically, increasing accuracy and reducing the burden on clinical staff.

Alzheimer's disease is the most common type of dementia, and other health problems can manifest in addition to cognitive impairment. The process of predicting the disease is time-consuming and in the early stages of Alzheimer's disease, mild cognitive impairment (MCI) can be diagnosed as a very mild manifestation of the disease. Speech impairment is one of the earliest symptoms of AD patients. Many studies have demonstrated the potential of automated acoustic assessment using acoustic and linguistic features extracted from speech.

This is an open access article under the CC BY-NC license
https://creativecommons.org/licenses/by-nc/4.0/

1 Analysis of Alzheimer's and Parkinson's Disease Prediction Based on Speech

In this section, we focus on significant advances and applications of machine learning in the detection of AD, MCI, and PD between 2000 and 2025, highlighting the best results achieved in the detection and prediction of these neurodegenerative diseases.

1.1 Analysis of Research for the Prediction of Alzheimer's Disease

Studies using automatic speech recognition (ASR) have focused on improving AD classification. The ADReSS (2020) and ADReSSo (2021) challenges evaluated speech-based cognitive decline detection [1, 2, 3]. ADReSS achieved 88% accuracy using only transcriptions without audio, and by combining audio with text features, accuracy increased to 93.8% [4]. Although manual transcriptions achieve the highest accuracy, ASR outputs were comparable to them [5]. A thoroughly tuned BERT model on ASR transcriptions achieved similar performance to manual transcriptions [6]. The study [7] compared three ASR systems (Google Speech, Whisper, and Wave2vec2), and Google achieved the highest accuracy of AUC 0.87.

The latest research on Alzheimer's disease and mild cognitive impairment [8] has focused more on the use of digital biomarkers for early diagnosis, reviewing 431 studies from five online databases. The research also focused on the use of automated speech analysis as a non-invasive biomarker for early diagnosis of AD in other languages. We focused mainly on studies with a larger number of AD cases. The study [9] analysed the acoustic characteristics of speech in 8779 Japanese seniors who read short sentences, and machine learning achieved a classification accuracy of AUC 0.61–0.77 in distinguishing MCI from healthy controls, with individuals with global cognitive impairment defined using the MMSE in the range of 20–23. In the publication [10], they applied a simple phonemic verbal fluency task in Thai to 100 individuals and proposed new feature extraction methods based on phonemic grouping and switching, achieving 65.5% accuracy in MCI classification.

Yamada et al. [11] compared speech responses to everyday questions in 94 Japanese participants with neuropsychological tasks and found that everyday questions can detect MCI with 87.4% accuracy. A study [12] of an NLP model based on simple picture description tasks in Slovak recorded using a smartphone achieved 95% accuracy in AD classification after balancing the data of healthy and AD individuals in a sample of 1117 people, of whom approximately 10% were diagnosed with AD and MCI. The study [13] used lexical-semantic and acoustic speech characteristics to identify MCI and AD in the early stages in participants from an English-speaking population in the US, with lexical-semantic scores achieving an AUC of 0.80 and correlating with amyloid-β biomarkers in cerebrospinal fluid. Ambrosini et al. [14] conducted a multilingual study in Italy and Spain with 133 participants, where automatic analysis of spontaneous speech enabled the classification of cognitive decline with 84% accuracy in participants with mild impairment. In their publication [15], they implemented a speech analysis algorithm in a Spanish-speaking population using five language tasks and machine learning, achieving an AUC of 0.93 and 88.4% accuracy in distinguishing MCI and dementia from healthy individuals.

1.2 Analysis of Research for the Prediction of Parkinson's Disease

Parkinson's disease, the second most common neurological disorder after AD, causes not only movement disorders but also speech disorders. Changes in speech, such as reduced volume, slurred speech, or changes in speed, can serve as early and non-invasive indicators of the disease.

Research in the field of PD diagnostics has shifted significantly towards the use of speech signals as a non-invasive biomarker. In a review study [16], the authors focused on the current state of research in the field of clinical decision-making using speech signal analysis in AD and PD. Based on selected clinical and technical criteria, an analysis of 72 selected studies on the latest developments in the field of early warning of PD based on speech was performed.

Zhang et al. [17] developed a mobile system for analysing linear and nonlinear dysphonia features, achieving 88.7% accuracy in a large Chinese cohort and deploying it via a mobile app, enabling practical telemedicine applications. In the US, a web-based PD screening framework with a short speech task using XGBoost model [18], reached an AUC of 0.753, with interpretable outputs via additive explanations. Vasquez-Correa et al. [19] integrated speech and movement signals into end-to-end deep learning, with models trained on smartphone and laboratory data achieving more than 92% accuracy. A Chinese study [20] using 3-second /a/ and /u/ vowels and a FastAI deep learning model reported 82% accuracy. Wang et al. [21] applied LASSO-based feature selection and ML classifiers, achieving 76% accuracy and highlighting articulation as more discriminative than phonation. A study [22] conducted as part of the EWA research project obtained data from a mobile application and balanced the originally unbalanced data set (117 PD vs. 944 HC), achieving the best PD detection result with an accuracy of 96.9%. An Italian study using SVM analysed voice across PD stages [23], identified clinical-instrumental correlations, quantified L-Dopa effects, and achieved 98.3% accuracy.

A Chinese study [24] using three PD datasets and advanced speech processing achieved up to 95% diagnostic accuracy, identifying key variable features. Hemmerling et al. [25] applied Vision Transformers to mel-spectrograms with explainable AI, reaching 89.8% accuracy and improving model interpretability. These findings confirm that combining high-quality voice data with advanced algorithms enables effective, accessible, and interpretable PD diagnostics.

2 Analysis of Data from the EWA-DB Database Using Machine Learning Methods

Phones and tablets are becoming increasingly important in predicting neurological diseases. As part of our collaboration on the **EWA project** [26] with AXON PRO and the Institute of Informatics of the Slovak Academy of Sciences in Bratislava, we used smartphones to record healthy and sick patients with AD, MCI, and PD. We created the EWA-DB database, from which we used data to analyse the detection of three classes: HC, AD-MCI, and PD.

2.1 EWA-DB Database for the Diagnosis of Alzheimer's and Parkinson's Disease

As part of the EWA project, the EWA-DB database was created to process speech parameter data and evaluate parameters to distinguish between the speech of healthy people and people with cognitive disorders. The database is supplemented and modified in line with the latest research, as is the resulting mobile application for smartphones [26]. EWA-DB is a Slovak database of voice recordings created for the purpose of researching the early diagnosis of neurodegenerative diseases, particularly Alzheimer's disease, mild cognitive impairment, and Parkinson's disease [27, 28]. The database contains various speech tasks: phonation of the vowel "a", diadochokinesis "pa-ta-ka", naming of objects and activities, and description of pictures.

The database also contains detailed demographic data, cognitive test results (e.g., MoCA), and information on meeting the inclusion criteria. Transcriptions were first automatically generated and then manually edited and annotated by trained annotators.

Annotations recorded various types of hesitations, unspoken sounds, unclear expressions, phonetic errors, and distracting sounds.

EWA-DB is publicly available through the ELDA platform and represents a significant contribution to the development of automated systems for the diagnosis of neurodegenerative diseases based on speech [27]. Due to its scope, quality, and linguistic specifics, it is suitable for research in the fields of artificial intelligence, linguistics, psychology, and clinical neurology.

2.2 Preprocessing and Feature Scaling

The extracted acoustic and linguistic features from EWA-DB are high-dimensional and exhibit non-uniform, heavy-tailed distributions. To ensure numerical stability of the models and reduce the influence of outliers, we applied **Quantile Transformer** scaling [29]. This transformation maps each feature to a target uniform or normal distribution, which reduces skewness and extreme values, thereby improving the convergence of linear classifiers.

The choice of Quantile Transformer was motivated by its demonstrated robustness in prior research [12, 22], where it consistently outperformed linear scaling approaches on imbalanced, speech-derived datasets. We also compare the quantile transformer with Min-Max and Max-Abs scalers. In our comparative experiments, alternative linear scalers such as Min-Max or Max-Abs resulted in lower performance, particularly in sensitivity and MCC. Therefore, Quantile Transformer was selected as the default scaling method throughout all models.

Principal component analysis (PCA) is a linear dimensionality reduction method that transforms the original variables into a new set of orthogonal components, ranked according to the amount of explained variability [30]. The algorithm involves standardising the data, calculating the covariance matrix, and finally projecting the data into a lower-dimensional space. PCA allows you to simplify the data structure, remove redundancy, and highlight the most important patterns in the data, which is especially useful in data preprocessing and classification.

2.3 Synthetic Minority Oversampling Technique

The dataset is strongly imbalanced, with healthy individuals substantially outnumbering patients with AD-MCI or PD. This imbalance poses a risk of bias in the classifier toward the majority class, which is why we used a similar approach to that used in studies [12, 22]. To address this issue, we employed Synthetic Minority Oversampling Technique (**SMOTE**) during the training phase [31]. SMOTE synthetically generates new samples of the minority classes (AD-MCI and PD) in feature space, which results in a more balanced class distribution.

By enriching the dataset with additional minority samples, this approach mitigates the risk of majority-class dominance and improves the model's ability to detect early neurodegenerative changes in at-risk participants.

2.4 Application of Machine Learning Algorithms to Speech Analysis from EWA-DB

Machine learning (ML) methods and algorithms are effective tools in the diagnosis and treatment of various neurodegenerative diseases. In the case of some of these diseases, it is possible to estimate the current condition of the patient by analysing speech signals using ML techniques. For this reason, it is essential to pay attention to data preparation and preprocessing, including the use of scaling and normalization techniques.

One of the primary objectives of this study is to design and implement preprocessing procedures tailored for speech-based diagnostic modelling. To this end, we employed the Stochastic Average Gradient Augmented algorithm (SAGA) that integrates the strengths of various gradient-based techniques [32].

SAGA is particularly well-suited for weakly convex problems, requiring no manual adjustments, and it automatically adapts to the convexity level of the objective function. Its fast linear convergence and scalability make it an effective choice for large and sparse datasets commonly encountered in machine learning applications.

2.5 Methodological Framework and Evaluation Metrics for Diagnostics

However, the success of these methods essentially depends on the quality of the input data processed by the selected algorithm. To solve complex diagnostic tasks in selected neurodegenerative diseases, such as Alzheimer's and Parkinson's disease, the research team EWA-RT has developed and verified a general methodological framework. It uses a system of support programs (in Python) with a modular, interactive, and adaptive structure that allows the assessment of the disease status based on input data (Fig.1). For a comprehensive and clinically meaningful evaluation of the diagnostic capabilities of the multidimensional class model, we used a set of performance metrics: accuracy, sensitivity, specificity, and F1 score, each of which reflects a different dimension of predictive reliability [33, 34]. The Matthews correlation coefficient (MCC) extends the binary form, allowing for the evaluation of multi-class classification. It remains robust even when classes are significantly imbalanced and is particularly suitable for tasks in biomedicine and bioinformatics, where multi-class predictions with uncertain data are common [35].

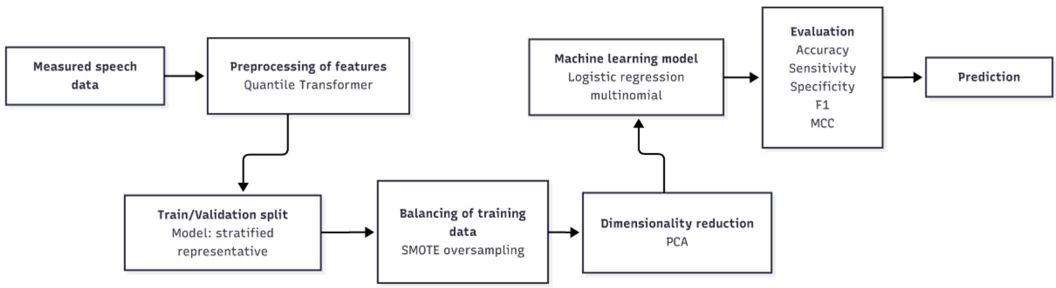


Fig.1. Block diagram of the ML modular system for detecting neurodegenerative diseases.

(Fig.1) presents a block diagram of the proposed machine learning pipeline for speech-based detection of Alzheimer’s and Parkinson’s diseases. The process begins with feature preprocessing and normalization using the Quantile Transformer to reduce noise and scale features. The dataset is then divided into training and validation subsets using a stratified representative split to preserve the class distribution. To address class imbalance, the SMOTE oversampling algorithm is applied exclusively to the training subset, ensuring that no synthetic samples are introduced into the validation data and thus preventing data leakage. Subsequently, dimensionality reduction is performed using Principal Component Analysis (PCA), followed by training a multinomial logistic regression model with SAGA optimization and L1 regularization. Model performance is evaluated using clinically relevant metrics, including Accuracy, Sensitivity, Specificity, F1-score, and the Matthews correlation coefficient (MCC). The final output represents a predictive model capable of classifying new speech recordings.

3 Classifying Neurodegenerative Diseases Based on Speech in Multiple Classes

To evaluate the effectiveness of machine learning for the early detection of neurodegenerative diseases from speech data, we developed and tested logistic regression–based classification model using the EWA-DB dataset. The dataset consists of high-dimensional acoustic and linguistic descriptors extracted from speech recordings, with participants labelled as healthy (HC), AD with mild cognitive impairment (AD-MCI), and Parkinson's disease (PD). Given the extreme dimensionality of the feature space (over 169 thousand features) and the inherent class imbalance (HC ≫ AD-MCI, PD), a standardized preprocessing pipeline was applied across both models to ensure comparability. Input features were scaled using a Quantile Transformer to normalize skewed distributions, and SMOTE oversampling was used to balance the minority classes (AD-MCI, PD) with the majority class HC. Dimensionality was then reduced with PCA, with thresholds between 82% and 95% of explained variance systematically evaluated. The model was trained using multinomial logistic regression with L1 regularization, optimized using the SAGA solver, which is suitable for high-dimensional problems with low density. The results shown in (Tab.3) showed the difference in training and testing accuracy as an additional indicator of stability to assess potential overfitting.

3.1 Stratified Selection of the Most Representative Samples

Our first idea was to build the validation set from the most representative cases; while keeping the more diverse/borderline cases for training so the model would learn robust decision boundaries. We therefore used a stratified split over two factors: diagnostic category (HC, AD-MCI, PD) and age bin (24–50, 51–60, 61–70, 71–80, 81–97). Each unique combination of diagnosis with age bin forms one stratum (e.g., “PD and 61–70”), (Tab.1), (Tab.2).

Within every stratum we measured the distance between samples in the scaled feature space and computed, for each sample, its average distance to all other samples in the same stratum. Samples with the smallest average distance (i.e., closest to the stratum centroid, thus most “typical”) were selected into the validation set; the remainder stayed in the training set. We selected 20% per stratum this way. The result is a validation set that faithfully reflects the population structure, and a training set enriched with more heterogeneous and boundary-case samples, which improves learning of discriminative class boundaries.

Table 1. Distribution of samples in the training and testing sets by diagnosis.

Dataset	HC	MCI-AD	PD	Total samples
Training set	823	96	129	1048
Validation set	204	24	31	259

Table 2. Distribution of samples in the training and testing sets by age bins.

Dataset	24-50	51-60	61-70	71-80	81-97	Total samples
Training set	25	290	398	243	92	1048
Validation set	7	73	98	59	22	259

The dataset was divided into a training set (1048 samples) and a validation set (259 samples). The distribution by diagnostic categories is summarized in (Tab.1), while the distribution by age bins is shown separately in (Tab.2). As shown in (Tab.1) and (Tab.2), the dataset was imbalanced, with healthy individuals forming the majority class. Most participants fell into the 51–70 age bins, while extreme bins (24–50 and 81–97 years) were sparsely represented. Preserving this demographic structure was essential, as speech features are strongly age-dependent and could bias classification if unevenly represented.

3.2 Machine Learning and Validation Workflow for Detecting AD-MCI and PD

To handle class imbalance, we applied SMOTE algorithm, which synthetically expanded the minority classes and yielded a balanced training set. Finally, we applied Principal Component Analysis (PCA) to reduce dimensionality. Out of the original more 169 thousand extracted features, PCA retained only 330 principal components while preserving 82% of the variance in the data (see (Tab.3). This dramatically reduced computational complexity and eliminated redundant or noisy features, while maintaining most of the clinically relevant information.

To further control overfitting in the high-dimensional feature space, the logistic regression model was trained with L1 regularization, which promotes sparsity by shrinking non-informative coefficients toward zero and thus performs embedded feature selection. For optimization we employed the SAGA algorithm, a stochastic variant of gradient descent that is particularly well-suited for large-scale and sparse problems. This combination enabled efficient training while focusing on the most discriminative speech features, further improving model generalization.

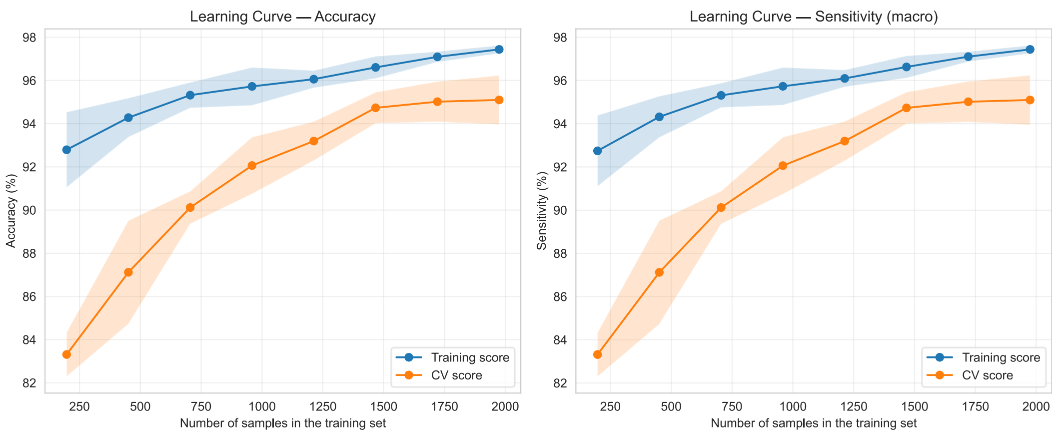


Fig.2. Cross-validation learning curves (accuracy and sensitivity).

To evaluate whether this training strategy leads to robust learning, we first inspected cross-validation learning curves (Fig.2). Both accuracy and macro-sensitivity increased steadily with larger training sets, while the gap between training and validation narrowed. This indicates that the model benefited from additional data and avoided severe overfitting, achieving good generalization capacity.

A complementary hold-out learning curve with the independent test set (Fig.3) further confirmed that test accuracy remained stable (approx. 92–93%) as the training size increased, supporting the stability and reproducibility of the model.

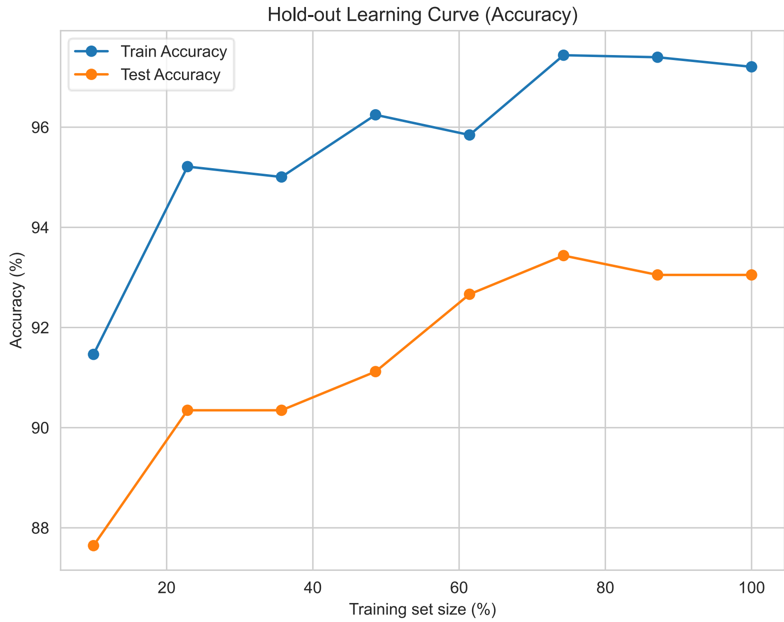


Fig.3. Hold-out learning curve on the independent test set.

The multinomial logistic regression model trained with this configuration achieved:

- Accuracy: 93.5%, Sensitivity: 93.6%, F1 Score: 88.5%
- MCC: 83.7%

(Fig.4) presents the confusion matrix for Model 1. The results demonstrate that the classifier performed consistently well across all three diagnostic categories.

- HC (class 0): The vast majority of samples (961) were correctly classified, with only a small proportion being misclassified.
- AD-MCI (class 1): Nearly all cases were correctly identified (115), with only 5 false negatives.
- PD (class 2): Most patients were classified correctly (146), with just a handful of errors.

Overall, the confusion matrix indicates a balanced and reliable performance across HC, AD-MCI, and PD participants. The strong diagonal dominance shows that the model not only detected the majority class accurately but also maintained high precision and recall for the minority classes, which are most critical in early neurodegenerative disease detection.

Receiver operating characteristic (ROC) analysis in (Fig.5) demonstrated high separability among all three classes. The area under the curve (AUC) was 0.989 for healthy controls, 0.996 for AD-MCI, and 0.988 for PD, with a macro-average AUC of 0.991. These findings further confirm that the classifier achieves both robust and clinically meaningful performance for the early detection of neurodegenerative disorders. The results indicate that representative stratified sampling, in combination with balanced training data and dimensionality reduction, enables reliable classification of neurodegenerative risk based on speech-derived features.

Confusion Matrix

		Predicted		
		HC	**AD**	**PD**
Actual	**HC**	961	33	33
	AD	5	115	0
	PD	6	8	146

Fig.4. The confusion matrix shows results for three classes: HC, AD-MCI, PD.

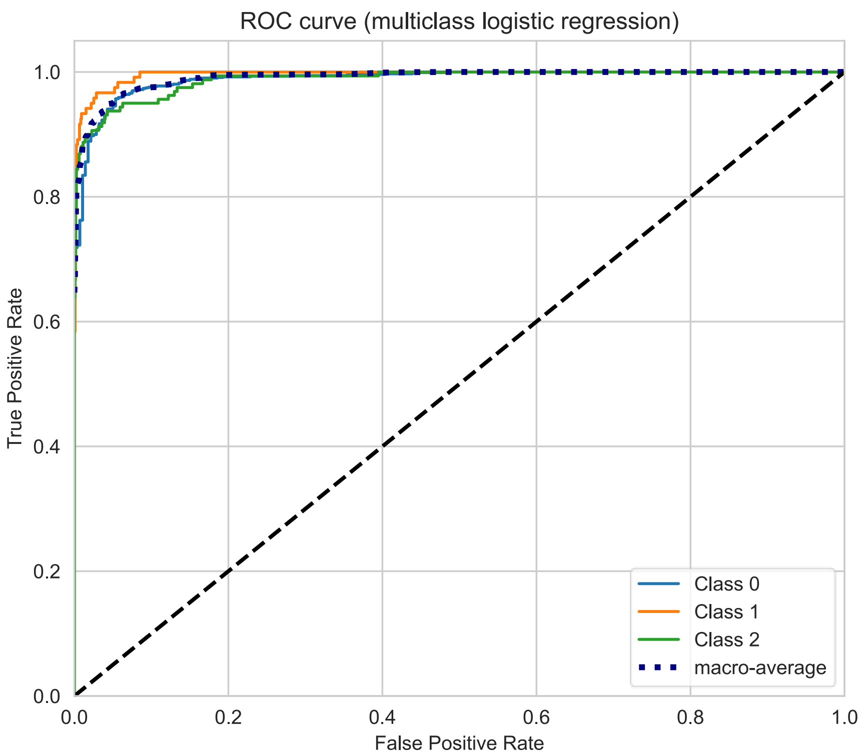


Fig.5. ROC curves for classes: Class 0 = HC, Class 1 = AD-MCI, Class 2 = PD.

3.3 Controlling Overfitting via PCA Variance Thresholds

To investigate the influence of dimensionality reduction on classification performance, we systematically varied the proportion of variance retained by PCA. Several thresholds were tested (70%, 75%, 82%, 85%, 88%, 90%, 92%, 95%), while keeping all other parameters constant (SMOTE oversampling, Quantile Transformer scaling, and logistic regression with L1 regularization).

For enhanced interpretability, (Tab.3) presents both the percentage of retained variance and the corresponding number of features (principal components) obtained after dimensionality reduction. This representation illustrates how the originally high-dimensional feature space was effectively compressed to a minimal set of features, while preserving most of the informative content in PCA %.

The primary evaluation criterion was performance on the independent test set, assessed by Accuracy, Sensitivity, F1 score, and MCC. We also considered the difference between training and testing accuracy, as this reflects the degree of overfitting or underfitting. The optimal configuration should therefore combine strong testing performance with a limited difference between training and testing accuracy.

As summarized in (Tab.3), the most advantageous Accuracy difference between training and testing (4.5 and 4.8, respectively) was observed at 82% and 85% PCA, as it optimizes the lowest risk of overfitting while maintaining competitive performance across all metrics. Although higher PCA thresholds (88%–92%) yielded slightly better F1 and MCC scores in the complete dataset, they also showed a larger difference between training and testing of more than 5%, indicating overfitting. Based on the balance between robust testing performance and model stability, PCA variance retention at 82% was chosen as the final model configuration.

Table 3. PCA is reported as retained variance (%)
and the resulting number of features and corresponding metrics in %.

PCA %	No. features	Accuracy Train-Test Gap	Sensitivity	F1 Score	MCC
70 %	175	3.5	90.3%	83.2	75.7
75 %	223	4.2	92.3	85.8	79.6
82 %	330	4.5	93.6	88.5	83.7
85 %	396	4.8	94.1	89.4	85
88 %	478	5.1	94.4	90.4	86.3
90 %	544	5.6	94.2	90.3	86.2
92 %	619	5.6	94.3	90.6	86.6
95 %	715	5.7	94.3	90.6	86.6

3.4 Alternative Sampling and Scaling Methods

In addition to evaluating different PCA thresholds, we systematically compared alternative data preprocessing strategies for the proposed model. The default pipeline combined SMOTE oversampling with Quantile Transformer scaling. To assess the robustness of this configuration, we substituted SMOTE with ADASYN for synthetic minority oversampling and replaced the Quantile Transformer with linear scaling methods (Min-Max and Max-Abs). All other pipeline components, including the logistic regression classifier and PCA-based dimensionality reduction, remained unchanged.

As summarized in (Tab.4), none of the alternative pipelines outperformed the default configuration. Specifically, ADASYN consistently resulted in lower sensitivity and MCC values, indicating reduced effectiveness in capturing minority classes. Similarly, Min-Max and Max-Abs scaling underperformed compared to the Quantile Transformer, underscoring the importance of nonlinear feature transformation for the highly skewed distributions present in the EWA-DB speech data. These findings confirm that the combination of SMOTE oversampling and Quantile Transformer scaling provides the most reliable preprocessing strategy for simultaneous AD-MCI and PD detection.

Table 4. Comparison of alternative sampling and scaling methods for the model.

Sampling	Scaling	Accuracy	Sensitivity	F1 Score	MCC
ADASYN	Min-Max	84.3%	85.4%	76.4%	65.3%
ADASYN	Max-Abs	83.9%	87.4%	76.4%	66.0%
SMOTE	Min-Max	83.1%	84.5%	74.9%	63.4%
SMOTE	Max-Abs	84.9%	85.4%	76.7%	66.3%

4 Discussion

In recent years, several solutions have been developed in the field of neurodegenerative disease detection, with the EWA-RT research team achieving comparable or higher accuracy than published studies. The discussion includes a comparison of selected results from the EWA-DB database with existing solutions, with the EWA-RT team's approach representing a new method of multi-class AD and PD detection based on linguistic and lexical characteristics from the EWA-DB database. The variability of speech tasks, extensive recordings, and detailed annotations provide a solid basis for evaluating algorithms and methodologies, enabling results comparable to international research.

The similar study from 2024 presented an automated speech analysis algorithm for detecting cognitive impairment (MCI and dementia) in Spanish-speaking individuals [15], Participants completed four short speech tasks, picture description ("Cookie theft"), phonemic fluency (words beginning with "F"), alternating semantic fluency (fruit and sports), and semantic fluency (animals), via a web or mobile application in less than five minutes. The recordings were processed using speech-to-text transcription and digital signal processing to extract acoustic and linguistic features. The final machine learning model achieved high classification performance: accuracy of 87.5–88.4%, sensitivity of 83.3–87.5%, and specificity of 89.2%.

Using the proposed methodology, the EWA-RT team achieved excellent classification results from the EWA-DB dataset, distinguishing patients with Alzheimer's disease, mild cognitive impairment (AD-MCI), and Parkinson's disease (PD) from healthy control subjects. The model achieved an accuracy of 93.3%, with a specificity of 93.5%, an F1 score of 88.4%, and a Matthews correlation coefficient of 83.2%. These results were enabled by rigorous data preprocessing, including stratified sampling by age and diagnosis, noise reduction and overfitting mitigation techniques, class balancing using SMOTE, and dimensionality reduction via principal component analysis while preserving key information.

Concurrently, multiple studies have examined the utility of the EWA-DB database for the prediction of Parkinson’s disease (PD). In 2022, the EWA-RT team demonstrated the effectiveness of machine learning in speech-based neurocognitive assessment of PD, achieving a classification accuracy of 96.9%, specificity of 89.3%, and Matthews correlation coefficient (MCC) of 83.2% [22]. The best results were achieved using linear regression in combination with Quantile Transformer scaling.

Further research explored cross-linguistic differences between Spanish and Slovak in the selection of acoustic features relevant to PD detection, utilizing data from the EWA-DB and PC-GITA datasets [36, 37]. The highest predictive performance within EWA-DB was obtained by combining optimized feature sets with spontaneous speech, resulting in an accuracy of 69.6% [36].

Additionally, the BDHPD study [37] introduced a novel deep learning architecture designed to enhance language generalization through self-supervised representation learning and transformation mechanisms. In its best single-language configuration, BDHPD achieved an accuracy of 83.6% and an F1 score of 70.3%.

Future work should focus on several directions to improve the EWA-DB framework. Extending the approach to preliminary classification by including Alzheimer's disease and mild cognitive impairment as separate classes alongside healthy controls would im-prove differentiation at an early stage. Similarly, distinguishing early-stage Parkinson's disease (PD) from HC is essential for early intervention. Combining linguistic characteristics with biomarkers such as neuroimaging, acoustic data, and cognitive scores could in-crease robustness and interpretability. The incorporation of longitudinal modelling to capture the temporal dynamics of disease progression (HC → MCI → AD or PD) would support predictive diagnostics and personalized treatment strategies.

Conclusion

Future research should focus on expanding the EWA-DB database with long-term records of healthy individuals and patients, which will significantly advance the prediction of non-invasive diagnostic methods. The automatic speech database creation system allows speech to be collected from individuals in their home environment using common devices such as smartphones or laptops. This approach supports the automatic processing of voice responses in any language, including Slovak, and significantly reduces the cost and time required to create specialized speech databases. Such processing will enable rapid and longitudinal analysis of clinical data, allowing for disease prediction, personalized approach, and therapy design, thereby minimizing adverse effects. Ultimately, clinical interpretability remains essential for real-world application. The development of explainable AI techniques that provide transparent and clinically meaningful insights will be key to building trust and integration into healthcare workflows.

Acknowledgement

This paper is supported by project GAAA/2023/6 "Basic research of new innovative methods using virtual reality for early detection of neurodegenerative diseases" and GAAA/2024/27 "Research of algorithms for process modelling, control and visualisation in domains of applied informatics", Grant Agency Academia Aurea (GAAA) https://www.gaaa.eu/.

Abbreviations

AD	Alzheimer's Disease
ADASYN	Adaptive Synthetic Sampling
ADReSS	Alzheimer's Dementia Recognition through Spontaneous Speech
ADReSSo	Alzheimer's Dementia Recognition through Spontaneous Speech only
AI	Artificial Intelligence
ASR	Automatic Speech Recognition
AUC	Area Under the Curve
BDHPD	Bilingual Dual-Head Parkinson's Disease Model
EWA	project: Early Warning of Alzheimer
EWA-DB	Early Warning of Alzheimer Database

EWAT	Faculty of Informatics: EWA - Team
F1	F1 Score (Harmonic Mean of Precision and Recall)
HC	Healthy Controls
LASSO	Least Absolute Shrinkage and Selection Operator
MCC	Matthews Correlation Coefficient
MCI	Mild Cognitive Impairment
ML	Machine Learning
MMSE	Mini-Mental State Examination
MoCA	Montreal Cognitive Assessment
PCA	Principal Component Analysis
PD	Parkinson's Disease
SAGA	Stochastic Average Gradient Augmented
SMOTE	Synthetic Minority Oversampling Technique
SVM	Support Vector Machine
Wave2vec2	Wav2Vec 2.0 (Self-Supervised Speech Representation Model)
XAI	Explainable Artificial Intelligence
XGBoost	Extreme Gradient Boosting

References

[1] Luz, S., Haider, F., de la Fuente Garcia, S., Fromm, D., MacWhinney, B.: Alzheimer's dementia recognition through spontaneous speech. Front. Comput. Sci. 2021, 3, 780169.

[2] Luz, S., Haider, F., de la Fuente, S., Fromm, D., MacWhinney, B.: Detecting cognitive decline using speech only: The ADReSSo challenge. arXiv 2021, arXiv:2104.09356.

[3] Ding, K., Chetty, M., Noori Hoshyar, A., Bhattacharya, T., Klein, B.: Speech-based detection of Alzheimer's disease: A survey of AI techniques, datasets and challenges. Artif. Intell. Rev. 2024, 57(12), 325.

[4] Martinc, M., Haider, F., Pollak, S., Luz, S.: Temporal integration of text transcripts and acoustic features for Alzheimer's diagnosis based on spontaneous speech. Front. Aging Neurosci. 2021, 13, 642647.

[5] Soroski, T., da Cunha Vasco, T., Newton-Mason, S., Granby, S., Lewis, C., Harisinghani, A., et al.: Evaluating web-based automatic transcription for Alzheimer speech data: Transcript comparison and machine learning analysis. JMIR Aging 2022, 5(3), e33460.

[6] Li, C., Cohen, T., Pakhomov, S.: The far side of failure: Investigating the impact of speech recognition errors on subsequent dementia classification. arXiv 2022, arXiv:2211.07430.

[7] Heitz, J., Schneider, G., Langer, N., Calzolari, N., Kan, M.Y., Hoste, V., et al.: The influence of automatic speech recognition on linguistic features and automatic Alzheimer's disease detection from spontaneous speech. In: Proceedings of the International Conference on Computational Linguistics, ELRA and ICCL: 2024, pp. 15955–15969.

[8] Qi, W., Zhu, X., Wang, B., et al.: Alzheimer's disease digital biomarkers multidimensional landscape and AI model scoping review. npj Digit. Med. 2025, 8, 366. https://doi.org/10.1038/s41746-025-01640-z.

[9] Nagumo, R., Zhang, Y., Ogawa, Y., Hosokawa, M., Abe, K., Ukeda, T., et al.: Automatic detection of cognitive impairments through acoustic analysis of speech. Curr. Alzheimer Res. 2020, 17(1), 60–68.

[10] Metarugcheep, S., Punyabukkana, P., Wanvarie, D., Hemrungrojn, S., Chunharas, C., Pratanwanich, P.N.: Selecting the most important features for predicting mild cognitive impairment from Thai verbal fluency assessments. Sensors 2022, 22(15), 5813. https://doi.org/10.3390/s22155813.

[11] Yamada, Y., Shinkawa, K., Nemoto, M., Ota, M., Nemoto, K., Arai, T.: Speech and language characteristics differentiate Alzheimer's disease and dementia with Lewy bodies. Alzheimer's Dement. Diagn. Assess. Dis. Monit. 2022, 14(1), e12364.

[12] Képešiová, Z., Ružický, E., Kozák, Š., Malaschitz, R., Zimmermann, A.: Application of advanced machine learning algorithms for early detection of mild cognitive impairment and Alzheimer's disease. In: Proceedings of the 2023 International Scientific Conference on Computer Science (COMSCI), IEEE: Sozopol, Bulgaria, 2023, pp. 1–5. https://doi.org/10.1109/COMSCI59259.2023.10315946.

[13] Hajjar, I., Okafor, M., Choi, J.D., Moore, E., Abrol, A., Calhoun, V.D., Goldstein, F.C.: Development of digital voice biomarkers and associations with cognition, cerebrospinal biomarkers, and neural representation in early Alzheimer's disease. Alzheimer's Dement. Diagn. Assess. Dis. Monit. 2023, 15(1), e12393.

[14] Ambrosini, E., Giangregorio, C., Lomurno, E., Moccia, S., Milis, M., Loizou, C., Ferrante, S.: Automatic spontaneous speech analysis for the detection of cognitive functional decline in older adults: Multilanguage cross-sectional study. JMIR Aging 2024, 7, e50537. https://doi.org/10.2196/50537.

[15] Kaser, A.N., Lacritz, L.H., Winiarski, H.R., Gabirondo, P., Schaffert, J., Coca, A.J., Cullum, C.M.: A novel speech analysis algorithm to detect cognitive impairment in a Spanish population. Front. Neurol. 2024, 15, 1342907. https://doi.org/10.3389/fneur.2024.1342907.

[16] De Silva, U., Madanian, S., Olsen, S., Templeton, J., Poellabauer, C., Schneider, S., Narayanan, A., Rubaiat, R.: Clinical decision support using speech signal analysis: Systematic scoping review of neurological disorders. J. Med. Internet Res. 2025, 27, e63004. https://doi.org/10.2196/63004.

[17] Zhang, L., Qu, Y., Jin, B., Jing, L., Gao, Z., Liang, Z.: An Intelligent Mobile-Enabled System for Diagnosing Parkinson Disease: Development and Validation of a Speech Impairment Detection System. JMIR Med Inform. 2020, 8(9), e18689. https://doi.org/10.2196/18689

[18] Rahman, W., Lee, S., Islam, M.S., Antony, V.N., Ratnu, H., Ali, M.R., Hoque, E.: Detecting Parkinson Disease Using a Web-Based Speech Task: Observational Study. J Med Internet Res. 2021, 23(10), e26305. https://doi.org/10.2196/26305

[19] Vasquez-Correa, J.C., Arias-Vergara, T., Klumpp, P., Pérez-Toro, P.A., Orozco-Arroyave, J.R., Nöth, E.: End-to-End Modeling of Speech and Gait from Patients with Parkinson's Disease: Comparison between High Quality vs. Smartphone Data. In: ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE: Toronto, Canada, 2021, pp. 7298–7302.

[20] Tandjung, M.D., Wu, J.C.M., Wang, J.C., Li, Y.H.: An Implementation of FastAI Tabular Learner Model for Parkinson's Disease Identification. In: Proceedings of the 2021 9th International Conference on Orange Technology (ICOT), IEEE: Tainan, Taiwan, 2021. https://doi.org/10.1109/ICOT54518.2021.9680650

[21] Wang, Y., Li, X., Chen, H., Zhang, J.: Automatic Detection of Parkinson's Disease Using Chinese Speech Signals and Feature Selection Algorithms. J Biomed Eng. 2022, 39(2), 145–158.

[22] Képešiová, Z., Kozák, Š., Ružický, E., Zimmermann, A., Malaschitz, R.: Comparative Analysis of Advanced Machine Learning Algorithms for Early Detection of Parkinson Disease. Proceedings 2022, Cybernetics & Informatics (K&I), Visegrád, Hungary, pp. 1-6. https://doi.org/10.1109/KI55792.2022.9925942

[23] Suppa, A., Costantini, G., Asci, F., Di Leo, P., Al-Wardat, M.S., Di Lazzaro, G., Saggio, G.: Voice in Parkinson's Disease: A Machine Learning Study. Front. Neurol. 2022, 13, 831428. https://doi.org/10.3389/fneur.2022.831428.

[24] Yuan, L., Liu, Y., Feng, H.-M.: Parkinson disease prediction using machine learning-based features from speech signal. Service Oriented Computing and Applications, 18(1), 2024, 101-107. doi:10.1007/s11761-023-00372-w.

[25] Hemmerling, D., Zakrzewski, M., Wodzinski, M., Dudek, M., Gaciarz, F., Wojcik-Pedziwiatr, M., Rumezhak, T.: Improving AI Interpretability for Multilingual Parkinson's Disease Classification through Voice Analysis. In AAAI Bridge Program on AI for Medicine and Healthcare, PMLR: Palo Alto, CA, USA, 2025, pp. 49–55.

[26] **Project EWA. Home.**
Available online: https://www.projektewa.sk/en/index.html (accessed on 24 August 2025).

[27] Rusko, M. et al.: EWA-DB Early Warning of Alzheimer speech database. https://catalog.elra.info/en-us/repository/browse/ELRA-S0489/, (2023).98. EWA-DB – Early Warning of Alzheimer speech database. https://zenodo.org/records/10952480, https://doi.org/10.5281/zenodo.10952480, 2024.

[28] Rusko, M., Sabo, R., Trnka, M.; Zimmermann, A., Malaschitz, R., Ružický, E. et al.: Slovak database of speech affected by neurodegenerative diseases. Sci Data 11, 1320, 2024. https://doi.org/10.1038/s41597-024-04171-6.

[29] Bolstad, B.M., Irizarry, R.A., Åstrand, M., Speed, T.P.: A Comparison of Normalization Methods for High Density Oligonucleotide Array Data Based on Variance and Bias. Bioinformatics 2003, 19, 185–193. https://doi.org/10.1093/bioinformatics/19.2.185.

[30] Jolliffe, I.T., Cadima, J.: Principal Component Analysis: A Review and Recent Developments. Philos. Trans. R. Soc. A 2016, 374, 20150202. https://doi.org/10.1098/rsta.2015.0202

[31] Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research 2002, 16, 321–357. https://doi.org/10.1613/jair.953

[32] Defazio, A., Bach, F., Lacoste-Julien, S.: SAGA: A Fast Incremental Gradient Method with Support for Non-Strongly Convex Composite Objectives. Advances in Neural Information Processing Systems, 2014.

[33] Sokolova, M., Lapalme, G.: A Systematic Analysis of Performance Measures for Classification Tasks. Inf. Process. Manag. 2009, 45, 427–437. https://doi.org/10.1016/j.ipm.2009.03.002.

[34] Takahashi, K., Yamamoto, K., Kuchiba, A., Koyama, T.: Confidence Interval for Micro-Averaged F1 and Macro-Averaged F1 Scores. Applied Intelligence 2022, 52, 4961–4972. https://doi.org/10.1007/s10489-021-02635-5.

[35] Gorodkin, J.: Comparing Two K-Category Assignments by a K-Category Correlation Coefficient. Computational Biology and Chemistry 2004, 28, 367–374. https://doi.org/10.1016/j.compbiolchem.2004.09.006.

[36] Rekdal, M.: Speech Analysis for Automatic Parkinson's Disease Detection: Feature, Data, and Language Analysis for Performance Optimization. Master’s Thesis, Norwegian University of Science and Technology (NTNU), Trondheim, Norway, 2024.

[37] La Quatra, M., Orozco-Arroyave, J.R., Siniscalchi, M.S.: Bilingual Dual-Head Deep Model for Parkinson’s Disease Detection from Speech. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2025, pp. 1–5. https://doi.org/10.1109/ICASSP49660.2025.10889445.

Authors

Mgr. Barbora Jurkovcová
AXON PRO Ltd. Bratislava, Slovakia
jurkovcova@axonpro.sk

Her research interests include Applied informatics, modeling, visualization, and applications in medicine.

Assoc. prof. RNDr. Eugen Ružický, PhD.
Faculty of Informatics, Pan-European University, Bratislava, Slovakia
eugen.ruzicky@paneurouni.com

His research interests include Applied informatics, system analysis, modeling, visualisation and applications in medicine.

prof. Ing. Štefan Kozák, PhD.
Faculty of Informatics,
Pan-European University in Bratislava, Slovakia
stefan.kozak@paneurouni.com
His research interests include system theory, linear and nonlinear control methods, numerical methods, algorithms and software for modeling, control, signal processing, IoT, and embedded intelligent systems applications in a wide range of industrial, healthcare, banking and service sectors.

RNDr. Ján Lacko, PhD.
Faculty of Informatics, Pan-European University, Bratislava, Slovakia
jan.lacko@paneurouni.com
His research interests include digitization of objects from the field of cultural heritage, healthcare, industry, urban planning. and their display by various techniques, including virtual and augmented reality.

RNDr. Alfréd Zimmermann
AXON PRO Ltd. Bratislava, Slovakia
zimmermann@axonpro.sk
He is owner and CEO of the company. Professional and scientific interests include artificial intelligence in general and natural language processing.

International Journal
Information Technology Applications (ITA)

Volume 14, Number 1, July 2025

BRIDGING THE LANGUAGE GAP: EVALUATING AND ENHANCING SLOVAK LANGUAGE SUPPORT IN LARGE LANGUAGE MODELS

Patrik Skovajsa

Abstract:

This study investigates the current level of Slovak-language support in large language models (LLMs) and proposed practical pathways toward high-quality, resource-efficient deployment. I benchmarked several state-of-the-art open-source and commercial LLMs on a newly created set of 100 Slovak questions covering grammar, semantics, style, slang, translation, and complex constructions. I evaluated the answers automatically with OpenAI GPT-4o-mini. Results show that Google Gemma 3 27 B achieves near parity with GPT-4o while running on a single high-end GPU, outperforming LLaMA 3.1 70 B by 27 percentage points in overall quality and cutting latency by a factor of four. My findings highlight Gemma 3 27 B as the best current trade-off for Slovak, while underscoring the strategic need for a dedicated Slovak LLM built on open resources.

Keywords:

Slovak language, large language models, language evaluation, Gemma 3, natural language processing.

Introduction

Slovak, like many typologically related languages, presents unique challenges when deploying large language models (LLMs) without task or language-specific fine-tuning. Current open-source models such as **Meta LLaMA 3.1** with 70 billion parameters or **Mistral-Large** demand substantial computational resources in both training and inference. A recent leap in model quality most noticeably in **Google's Gemma 3** family, and especially the 27 billion-parameter varian suggests that these obstacles can be mitigated. Gemma 3 supports more than 140 languages and was explicitly designed for efficient single-machine deployment.

Multilingual encoder–decoder architectures like **mT5** already cover up to 101 languages, including Slovak. Without any additional adaptation they perform competitively in downstream tasks such as text classification, structured prediction, and open-domain question answering [1].

For the Slavic language group there exist models that have been tuned explicitly, for example **XLM-R**, which achieves state-of-the-art results in named-entity recognition, normalization, and entity linking across Czech, Polish, and Russian. With F_1 scores reaching 0.914 for Czech, these results indicate a strong likelihood of comparable success on Slovak data sets [2].

A fully Slovak-centric approach is represented by **SlovakBERT**, a RoBERTa-style transformer trained on a large Slovak web corpus. SlovakBERT attains excellent performance in morpho-syntactic tagging, sentiment analysis, document classification, and semantic textual similarity, and therefore constitutes a valuable building block for the Slovak NLP community [3].

Broader multilingual models such as **mGPT**, trained on Wikipedia and the C4 corpus in 61 languages, have demonstrated credible zero-shot performance in low-resource settings.

This is an open access article under the CC BY-NC license
https://creativecommons.org/licenses/by-nc/4.0/

Although Slovak was not a primary training target, the model's cross-lingual capabilities make it attractive for general NLP tasks that do not warrant Slovak-specific fine-tuning [4].

Among the models directly optimised for Slovak, **mistral-sk-7** a fine-tuned derivative of Mistral-7B trained on the Araneum Slovacum VII Maximum corpu provides a solid foundation for further customisation [5]. Iteratively fine-tuning such medium-sized backbones, including Mistral and Gemma, was emerged as a pragmatic route toward high-quality Slovak LLMs.

Progress on evaluation resources has kept pace. **SK-QuAD**, the first manually curated Slovak question-answering data set, contains over 91 000 factoid questions aligned with the SQuAD v2.0 format. By providing unanswerable questions and "plausible distractor" answers in addition to positive examples, SK-QuAD significantly improves zero-shot accuracy of multilingual models on Slovak QA tasks [6].

Complementary language-specific optimisations are equally important. The Slovak Morphological Tokeniser (**SKMT**), a byte-pair-encoding (BPE) variant that preserves stem integrity, can markedly boost downstream performance [7].

Like many non-dominant language communities, Slovakia faces systematic barriers to adopting state-of-the-art AI. Model training data rarely include sufficient Slovak text, and the few closed models that do support Slovak are unavailable for security-sensitive or confidential deployments. As a result, public institutions and companies struggle to exploit LLMs fully.

The European Union has begun to address this gap by creating funding frameworks for national-language models. Yet, Slovak development is complicated by the prevalence of Czech in shared corpora; close linguistic proximity often leads LLMs to conflate the two languages unless explicitly disambiguated. This phenomenon is common across closely related languages and underscores the need for models that can make finer distinctions.

Even where open-source models nominally support Slovak, practical use may be impossible: LLaMA 3.1-70B, for example, requires more than 200 GB of GPU memory merely to run inference.

Finally, conventional automatic metrics such as **BLEU** and **ROUGE** designed around English—poorly capture quality in morphologically rich languages like Slovak. Current research therefore calls for language-sensitive evaluation methods [8]. In this work I adopt OpenAI GPT-4o-mini, a proprietary model with high Slovak proficiency, as an automatic judge to benchmark various open-source LLMs.

The following section first outlines the experimental design used to evaluate Slovak-language competence across selected LLMs, including the newly created 100-question benchmark and the six-component evaluation scale (grammar, semantics, style & context, slang and regional expressions, translation quality, and complex constructions) together with an automated scoring pipeline built on GPT-4o-mini.

The **Evaluating LLM Performance in Slovak** block summarizes practical insights from the test runs, detailing the custom Python helper that ensured consistent, reproducible scoring. Finally, **Results and Model Comparison** synthesizes the quantitative findings, highlights the qualitative leap delivered by Google's Gemma3:27B, and discusses the trade-offs between cloud-based and on-premise deployment, pointing to Gemma3:27B as the current best balance of Slovak accuracy and hardware efficiency.

1 Evaluating LLM Performance in Slovak

As shown in (Fig.1), generating a single answer with the LLaMA-3.1-70B model can take nearly five minutes. In a full run of 100 diverse Slovak questions, the evaluation was never completed without errors; one attempt lasted more than six hours. This makes the model impractical for large-scale evaluations.

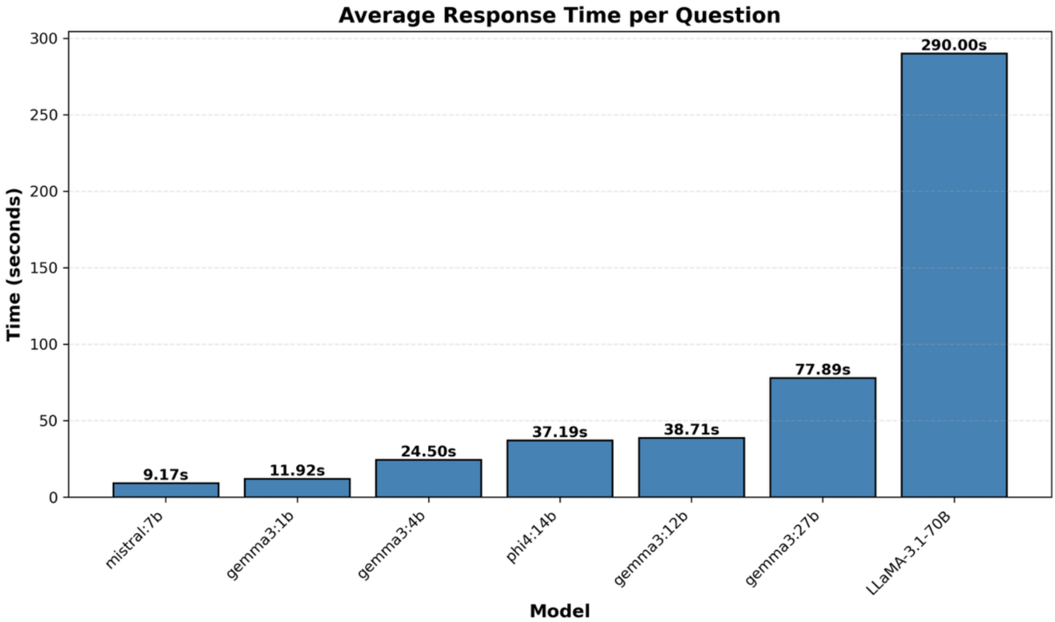


Fig.1. Average response time per question for each evaluated model.

For an in-depth analysis, I posed 100 diverse Slovak-language questions to several language models, assessing each response against six clearly defined criteria:

1. **Grammar**—evaluating the grammatical correctness of the generated sentences.
2. **Semantics**—assessing the accuracy of the conveyed meaning.
3. **Style and context**—determining the model's ability to maintain contextual relevance and appropriately answer follow-up queries.
4. **Slang and regional expressions**—measuring the understanding and accurate usage of Slovak slang and regional dialects.
5. **Translation**—examining the translation quality between Slovak and other languages.
6. **Complex constructions**—testing the models' handling of advanced grammatical structures and nuanced linguistic expressions.

To automate and streamline the evaluation process, an orchestration workflow was implemented (Fig.2). This system systematically submits test questions, collects model responses, and uses the GPT-4o-mini judging service to score them across six criteria, ensuring accuracy and consistency in the evaluation.

2 Results and Model Comparison

The evaluation results are presented in (Tab.1). They reveal a significant qualitative leap in the latest language models, particularly in Gemma 3:27B. Google's Gemma 3 series, available in 1B, 4B, 12B, and 27B parameter versions, features multimodal capabilities and context windows reaching up to 128k tokens. These models offer extensive language support exceeding 140 languages and are specifically engineered for efficient deployment on resource-limited systems, making genuine edge computing scenarios feasible.

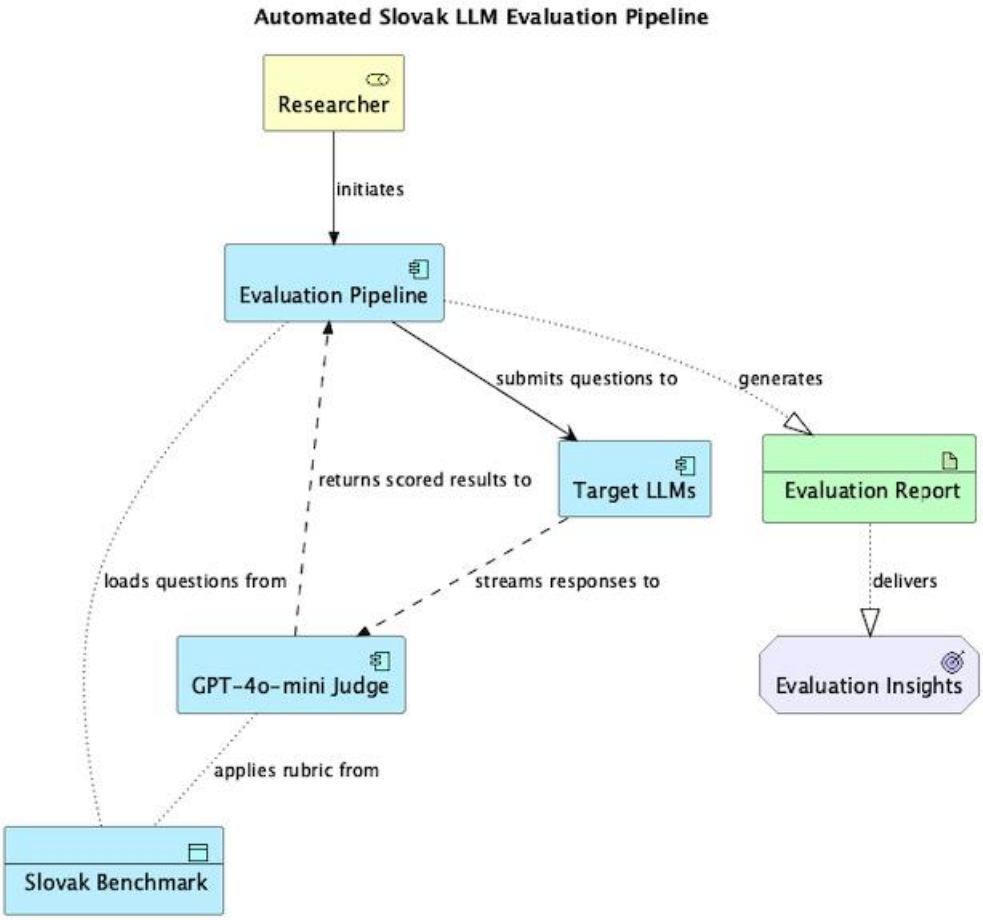


Fig.2. Automated Slovak LLM evaluation workflow.

Table 1. Evaluation results and response times of selected large-language models.

Large Language Model	Response time (s)	Grammatical accuracy	Semantic precision	Context retention	Understanding slang and regional expressions	Translation quality	Handling complex language constructions
gemma3:27b	77.89	9.1	9.2	9.9	8.2	8.1	8.8
gemma3:12b	38.71	8.9	9.0	9.8	8.0	7.5	8.5
gemma3:4b	24.5	8.0	8.2	9.0	7.0	6.2	7.6
deepseek-r1:32b	81.75	7.4	8.2	8.4	6.4	5.5	7.5
phi4:14b	37.19	7.0	7.6	8.0	5.8	5.8	6.8
deepseek-r1:14b	37.35	6.0	6.8	6.8	4.8	3.8	6.0
gemma3:1b	11.92	4.7	4.9	5.8	3.6	3.4	4.6
granite3.2:8b	24.22	4.7	5.6	6.0	4.2	3.5	5.0
mistral:7b	9.17	4.4	5.1	5.6	4.0	3.1	4.5

Identifying an appropriate language model for Slovak remains challenging due to limited native-language support, significantly hindering regional AI progress. Additionally, operational expenses, especially the stark contrast between cloud-based and on-premise deployments, present substantial hurdles.

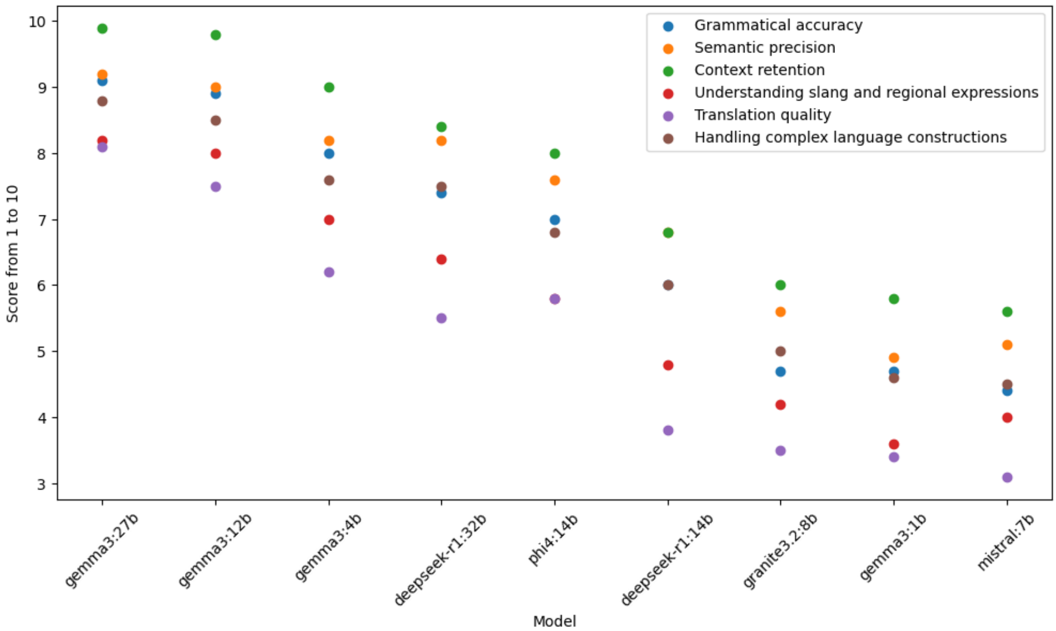


Fig.3. Performance comparison of large language models by efficiency and accuracy.

Several factors influence the choice between cloud and on-premise deployments. Cloud solutions offer scalability, reduced initial investment, and flexible cost structures. Conversely, on-premise deployments necessitate substantial initial expenditures on hardware but offer superior control over data privacy and security. Organizations must strategically evaluate these factors based on their specific operational requirements and resources.

Initially, models from OpenAI such as GPT-3 and subsequently GPT-4o established a robust benchmark for Slovak-language support. Google's Gemini later enhanced this landscape. Nevertheless, there remained a notable gap for a robust, open-source solution suitable for efficient on-premise utilization. Their relative efficiency and accuracy are visualized in (Fig.3).

Experiments with open-source alternatives, notably Meta's LLaMA 3.1-70B, yielded underwhelming results. Although Slovak was nominally supported, the quality of responses was inadequate, coupled with impractical hardware demands. Subsequent generations marked substantial improvements, notably the Mistral-Large, which significantly enhanced Slovak handling and achieved commendable performance.

The real breakthrough emerged with Gemma 3:27B, whose Slovak language proficiency closely rivals that of GPT-4o while remaining practically viable for on-premise deployment scenarios. Effective native-language support extends beyond mere convenience, becoming a critical component of national security, given its profound impact on the capability to directly analyze and manage local data, thereby reinforcing cyber-resilience.

3 Related Work

Ahuja et al. [9] introduced MEGAVERSE, a multilingual benchmark covering 22 datasets and 83 languages, many of them low-resource [9]. While MEGAVERSE provides broad cross-lingual comparisons, it does not address whether mid-sized open models can reach near state-of-the-art performance in specific languages. The evaluation reported here shows that, for Slovak, the Gemma 3:27B model achieves results comparable to proprietary models, suggesting that efficient alternatives exist even in constrained settings.

Skadiņa et al. [10] constructed the first Latvian benchmark by translating tasks such as COPA and MMLU, and demonstrated that manual post-editing significantly improves model accuracy. The study reported here avoids translation artifacts by employing a manually crafted Slovak dataset, highlighting the importance of high-quality, native-language resources in morphologically rich contexts.

Ojo et al. [11] proposed AfroBench, a benchmark evaluating 64 African languages across 15 tasks, and reported substantial performance gaps between proprietary and open-source models [11]. A similar disparity was confirmed for Slovak through the experimental results presented in this work. However, the evaluation also demonstrates that optimized tokenization and targeted data curation can substantially reduce this performance gap.

Arnett & Bergen [12] investigated the underperformance of LLMs in morphologically complex languages and attributed the issue primarily to reduced effective training data rather than linguistic structure itself. This interpretation is supported by the results presented here, which indicate that increasing the availability of Slovak-language data allows models like Gemma3:27B to nearly match English-language performance.

Azime et al. [13] in the ProverbEval benchmark, highlighted that culturally grounded tasks such as proverb understanding tend to increase performance variability across languages. The study reported here responds to that challenge by incorporating "slang and regional expressions" as one of six evaluation categories, exposing limitations even in otherwise strong models.

Conclusion

The evaluation of large language models (LLMs) for Slovak has highlighted significant advancements, particularly with Google's Gemma 3:27B model. This model demonstrates a remarkable balance between performance and efficiency, making it a viable option for on-premise applications. Its ability to handle complex linguistic structures and provide accurate translations underscores its potential in various domains.

However, relying solely on foreign-developed models poses challenges. While models like Gemma 3:27B offer impressive capabilities, they may not fully capture the nuances of the Slovak language or cater to specific national needs. Moreover, dependence on external models raises concerns about data sovereignty, security, and long-term sustainability.

Although (Tab.1) shows that Gemma 3:27B scores above 9 in grammatical accuracy and semantic precision—and maintains excellent context retention—a closer look at its output still reveals several micro-level deficiencies. The most common are orthographic mistakes (missing diacritics or wrong dash characters), morphological slips (incorrect adjective gradation or case forms), lexical/terminological inaccuracies (e.g., using *"časový základ - time base"* instead of the standard grammatical term *"časovanie – timing"*), inconsistent capitalisation of names, minor typographic issues, and occasional stylistic redundancy caused by overloaded bullet lists.

These errors stem from limited exposure to expertly proof-read Slovak data in the training corpus and from the language's highly inflected morphology, which introduces many exceptions the model only partly "guesses" correctly.

For most everyday use cases the text remains fluent and perfectly comprehensible—especially when contrasted with older open-source LLMs such as mistral-7b, whose grammatical and semantic scores hover around 4–5 in (Tab.1). However, for official publications or legal documents a human proof-reader is still indispensable to eliminate the systematic fine-grained errors current models cannot yet fully capture.

To address these issues, it is imperative for Slovakia to invest in developing its own LLMs. Such an initiative would not only enhance the country's technological autonomy but also ensure that the unique characteristics of the Slovak language are adequately represented and preserved. By building a dedicated Slovak LLM, tailored to the nation's linguistic and cultural context, Slovakia can foster innovation, support local industries, and strengthen its position in the global AI landscape.

In conclusion, while leveraging existing models like Gemma 3:27B is beneficial in the short term, the strategic development of a national Slovak LLM is essential for long-term growth, resilience, and cultural preservation. This endeavor will empower Slovakia to harness the full potential of AI technologies while safeguarding its linguistic heritage.

References

[1] Xue, L., Constant, N., Roberts, A., Kale, M., Al-Rfou, R., Siddhant, A., Barua, A., Raffel C., (2021). Google Research, Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.

[2] Viksna, R., Skadin, I., Deksne, D., Rozis, R., (2023). Large Language Models for Multilingual Slavic Named Entity Linking, Proceedings of the 9th Workshop on Slavic Natural Language Processing 2023, Association for Computational Linguistics.

[3] Pikuliak, M., Grivalský, Š., Konôpka, M., Blšták, M., Tamajka, M., Bachratý V., Šimko, M., (2022). SlovakBERT: Slovak Masked Language Model. In Findings of the Association for Computational Linguistics: EMNLP 2022 (proceedings). Abu Dhabi, United Arab Emirates.

[4] Shliazhko, O., Fenogenova, A., Tikhonova, M., Kozlova, A., Mikhailov, V., Shavrina, T., (2025). mGPT: Few-Shot Learners Go Multilingual. Institute of Linguistics RAS, Russia.

[5] Bednár P., Dobeš, M., Garabík R., (2023). Mistral-sk-7b. Hugging Face. Štúr Institute of Linguistics, Slovak Academy of Sciences, supported by DiusAI a. s..

[6] Hládek, D., Staš, J., Juhár J., Koctúr O., (2023). Slovak Dataset for Multilingual Question Answering.

[7] Držík, D., Forgáč, F., (2024). Slovak morphological tokenizer using the Byte-Pair Encoding algorithm.

[8] Dobeš, M., (2025). Evaluation of quality of Slovak language use in LLMS, Acta Electrotechnica et Informatica, Vol. 25, No. 1, 2025.

[9] Ahuja, S., Aggarwal, D., Gumma, V., Watts, I., Sathe, A., Ochieng, M., Sitaram, S. (2024). MEGAVERSE: Benchmarking Large Language Models Across Languages, Modalities, Models and Tasks. Proceedings of NAACL-HLT 2024. Mexico City, Mexico.

[10] Skadiņa, I., Bakanovs, B., Darģis, R. (2025). First Steps in Benchmarking Latvian in Large Language Models. Proceedings of RESOURCEFUL-2025. Tallinn, Estonia.

[11] Ojo, J., Ogundepo, O., Oladipo, A., et al. (2023). AfroBench: How Good Are Large Language Models on African Languages.

[12] Arnett, C., Bergen, B. (2025). Why Do Language Models Perform Worse for Morphologically Complex Languages. Proceedings of COLING 2025.

[13] Azime, I. A., Tonja, A. L., Belay, T. D., et al. (2025). ProverbEval: Exploring LLM Evaluation Challenges for Low-resource Language Understanding. In Findings of the Association for Computational Linguistics: NAACL 2025 (proceedings). Albuquerque, New Mexico.

Authors

Bc. Patrik Skovajsa
Pan-European University,
Faculty of Informatics, Bratislava, Slovakia
patrik.skovajsa@gmail.com
Secure AI solutions, risk management, LLMs, IT Expert, build the future with AI, one innovation at a time—ready to lead or collaborate on groundbreaking projects.

APPLICATIONS OF REINFORCEMENT LEARNING IN MODELING AUTONOMOUS BEHAVIOUR

Branislav Valacsay

Abstract:

This study presents a practical application of reinforcement learning (RL) for autonomous vehicle control within the Gymnasium CarRacing-v0 simulation environment. The primary objective was to design and train an agent capable of autonomously navigating procedurally generated racetracks by employing deep reinforcement learning methodologies.

The research utilized the Proximal Policy Optimization (PPO) algorithm, implemented via the Stable-Baselines3 library, chosen for its proven efficiency and robustness in continuous action domains. The training process was conducted in Python, using the PyTorch and Gymnasium frameworks.

Particular emphasis was placed on the agent's ability to develop smooth steering and throttle control—capabilities essential for handling sharp curves and maintaining optimal velocity. The findings demonstrate that a properly configured PPO agent can successfully learn to navigate the CarRacing environment after sufficient training. The study concludes with a performance evaluation, an analysis of the agent's acquired behavior, and suggestions for future research avenues.

Keywords: *Reinforcement learning, PPO, Gymnasium, CarRacing-v0, autonomous agent, continuous control.*

Introduction

Autonomous vehicle control represents one of the most significant challenges in contemporary artificial intelligence, requiring the effective integration of advanced machine learning algorithms with realistic simulation environments.

Among the most promising approaches to addressing such tasks is reinforcement learning (RL), which enables an agent to acquire optimal behaviour based on feedback received from the environment. Due to increasing computational power, high-quality libraries, and publicly available simulation platforms, reinforcement learning is becoming progressively more accessible to students and researchers without the need for extensive infrastructure.

The motivation for pursuing this topic stems from the author's personal interest in gaining practical experience with the application of reinforcement learning in control tasks, within the context of the Student Scientific Conference (ŠVOČ). At the same time, the chosen topic serves as a preparatory foundation for a subsequent bachelor's thesis. The aim of this contribution is to design, implement, and train an autonomous agent capable of controlling a virtual vehicle within the CarRacing-v0 environment, which is part of the Gymnasium library. This environment features procedurally generated racetracks and a continuous action space, making it a challenging benchmark for adaptive control algorithms.

To address the task, the Proximal Policy Optimization (PPO) algorithm was selected, known for its stability and performance in continuous action domains. The training process was conducted using the Stable-Baselines3, PyTorch, and Gymnasium libraries, with various learning parameters tested to achieve the highest possible driving performance of the agent. The results indicate that with a sufficient number of training episodes, the agent can execute smooth turns and maintaining speed without veering off track.

This is an open access article under the CC BY-NC license
https://creativecommons.org/licenses/by-nc/4.0/

The following chapters present an overview of the current state of knowledge in reinforcement learning for simulated control, followed by a description of the research design, including the tools and experimental parameters used. The discussion focuses on evaluating the agent's performance and identifying the strengths and limitations of the adopted methodology. The paper concludes with a summary of key findings and suggestions for potential future improvements.

1 State of the Art

Reinforcement learning (RL) is a subfield of machine learning that focuses on training an agent through interaction with its environment, based on feedback in the form of rewards and penalties. This approach is particularly well-suited for tasks that require sequential decision-making and long-term return optimization. A typical example of such a task is autonomous vehicle control, where the agent must coordinate steering, acceleration, and breaking in response to visual and dynamic stimuli.

One of the early breakthrough successes of RL in the domain of games and simulations was the application of the Deep Q-Networks (DQN) algorithm, which combines classical Q-learning with deep neural networks. DQN achieved significant results in discrete action spaces, such as those found in Atari games.

However, its applicability in environments requiring continuous control is considerably limited. The discreteness of actions prevents fine-tuned modulation of control signals, which is essential for maintaining speed and negotiating curves smoothly, as required by the CarRacing-v0 environment.

In this context, the Proximal Policy Optimization (PPO) algorithm has proven to be particularly effective. PPO belongs to the class of on-policy methods and is characterized by stable learning, even in the presence of complex visual inputs.

It employs a clipping mechanism on the probability ratio between the old and new policy to regulate weight updates, thereby reducing the risk of destabilizing the model during training. These properties make PPO a suitable choice for high-complexity environments with continuous action spaces. Owing to these characteristics, PPO has gained widespread adoption in visual RL tasks, including navigation, robotic simulations, and autonomous driving.

The CarRacing-v0 simulation environment, part of the Gymnasium library (the successor to OpenAI Gym), serves as a standard benchmark for evaluating the performance of RL algorithms in control tasks. Its variability, visual input in the form of RGB frames, and continuous action interface create realistic and challenging scenarios that thoroughly test an agent's ability to learn an optimal trajectory.

2 Design of Research

a) Theoretical Background

In reinforcement learning (RL), an agent is an autonomous computational entity that learns to perform actions within a given environment in order to maximize a numerical reward signal. The environment provides observations and rewards in response to the agent's actions, forming a closed-loop system known as the agent-environment interaction cycle.

At each time step, the agent receives an observation sts_tst, selects an action ata_tat, and receives a reward rtr_trt along with a new state s_{t+1}s_{t+1}s_{t+1}. The agent's objective is to learn a policy π(a|s)\pi(a \mid s)π(a|s) that maximizes the expected cumulative reward over time [1].

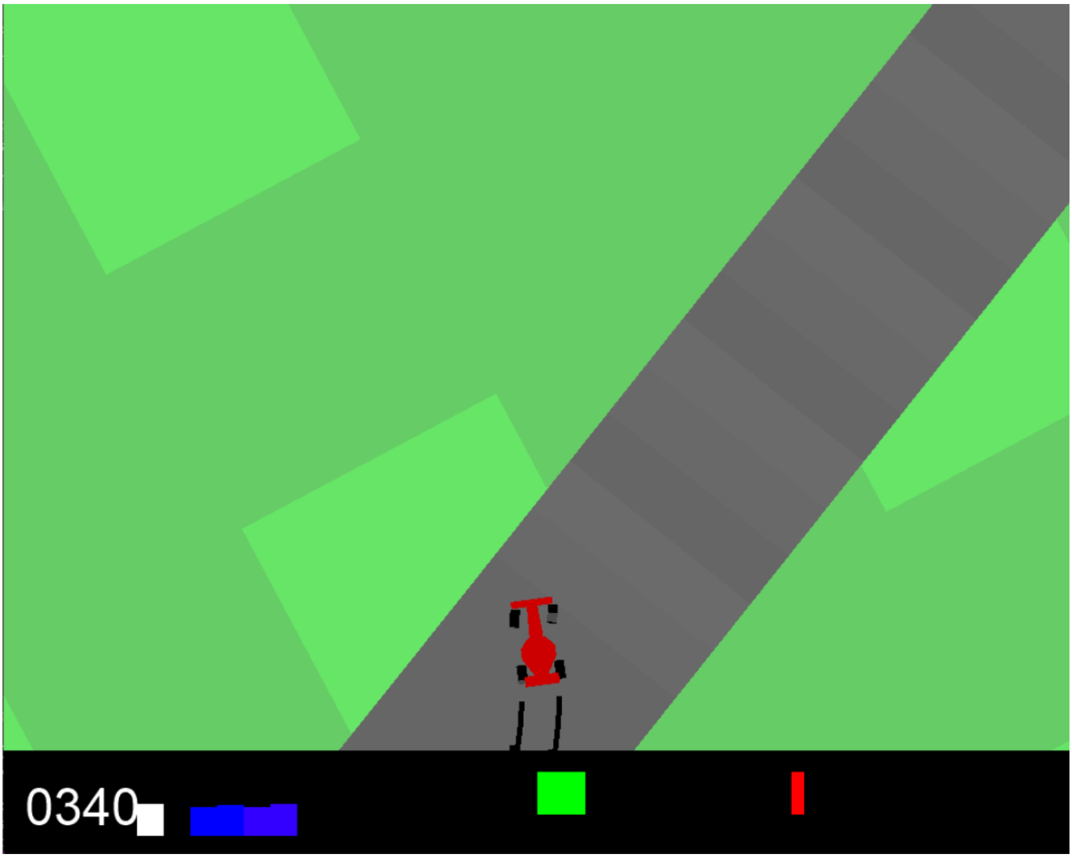


Fig.1. In game picture of agent actually playing the game.
Agent is capable of directly observe game screen and make decisions.

The learning process can be implemented through various RL algorithms. This work focuses on Proximal Policy Optimization (PPO), a state-of-the-art on-policy reinforcement learning method introduced by OpenAI. PPO belongs to the class of policy gradient methods, in which the agent directly optimizes a parameterized policy—typically represented by a neural network—by estimating the gradient of the expected reward.

A key innovation of PPO lies in its clipped surrogate objective function, which constrains large updates to the policy. This mechanism enhances training stability and mitigates the risk of destructive policy shifts.

The objective function used by PPO is defined as follows [1]:

$$L^{CLIP}(\theta) = E_t[\min(r_t(\theta)A_t, clip(r_t(\theta), 1-\in, 1+\in)A_t]$$

Where:

$$r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}$$

- denotes the probability ratio between the new and old policy, A_t is the advantage estimate at time step t, and ∈ is a hyperparameter that determines the range of the clipping, used to limit policy updates and enhance training stability.

Compared to traditional or older policy gradient methods (A2C – advantage actor-critic) or off policy methods such as Q-learning variants such as Deep Q-Networks (DQN), PPO offers a favorable trade-off between performance and implementation simplicity.

b) Object of Research

The object of this research is the development and training of an autonomous agent capable of efficiently controlling a vehicle within the CarRacing-v0 simulated environment (Fig.1). This environment, available through the Gymnasium library, represents a complex benchmark with procedurally generated tracks, visual input in the form of image frames, and a continuous action space, making it a suitable tool for evaluating the capabilities of reinforcement learning in tasks requiring smooth control.

c) Goal

The goal of the research is to implement and train an agent using the Proximal Policy Optimization (PPO) algorithm, which will be able to complete the track autonomously without leaving the road, with an emphasis on steering smoothness, adaptability to changing track layouts, and training stability.

d) Hypothesis

It is hypothesized that the PPO algorithm, applied within a continuous action space, can enable the agent to learn stable behaviour in the CarRacing-v0 environment, and that the resulting policy will be capable of generating smooth and effective control signals. Furthermore, it is assumed that a significant number of training episodes will be required for the agent to master the task due to the high visual and action-space resolution of the environment.

e) Methods of Validation

The training process was carried out in Python using the Stable-Baselines3 and PyTorch libraries. A PPO agent with a convolutional neural network architecture was employed, receiving 96×96 RGB image frames as input and producing three action values: steering, acceleration, and braking. The environment provided feedback in the form of a reward signal, with the objective of maximizing the cumulative reward per episode. Performance was validated by evaluating the average reward over the last 100 episodes and by observing the agent’s behaviour on unseen tracks.

f) Model, Data, and Configuration

The model consisted of two convolutional layers followed by fully connected layers, with the output layer representing the parameterized policy and the value function estimate. The data were generated on-policy by the simulator during training. Training was conducted in batches of 128 steps, using a learning rate of 3e-4, entropy penalty of 0.01, and a clipping factor of 0.2.

The Adam optimizer was used for optimization as implicit optimizer internally processed by Stable Baselines 3. The training process involved 5 000 000-time steps (Tab.1) and was executed on an NVIDIA RTX-series GPU.

Table 1. Training information

Parameter	Value
Total Timesteps	5 000 000
Training Duration	559 minutes 22 seconds
Environments	4 parallel
Algorithm	PPO
Average Reward	500+

g) Action Space

The observation-action space of the utilized game environment was intentionally converted to discrete finite set to facilitate action selection and enable more efficient agent learning.

h) Evaluation

The model's performance has been evaluated using the following metrics:

- average reward per episode,
- the agent's ability to complete a track,
- smoothness of driving and number of off-track events.

Additionally, the agent's behaviour on various procedurally generated tracks was visually analysed, offering a qualitative insight into the learned control strategies.

After sufficient training, the agent was able to maintain high speeds and consistently complete tracks without erratic movements, indicating the effectiveness of the applied algorithm in this environment.

3 Discussion

The experimental results confirm that the Proximal Policy Optimization (PPO) algorithm is a suitable choice for addressing the task of autonomous vehicle control in the CarRacing-v0 environment.

Following extensive training, the agent was capable of maintaining the vehicle on track, adapting to procedurally generated road curvatures, and ensuring smooth control without abrupt fluctuations in actions.

A key benefit of employing PPO was its ability to effectively handle the continuous action space, which is essential for realistic vehicle control—particularly for steering through curves and modulating speed.

The main advantages of the proposed solution include:

- **Training stability** in comparison with alternative algorithms such as DQN, which require discretization of the action space and thereby limit the agent´s ability to perform fine-grained control.
- **Practical implementation** within an environment accessible to a broad community of developers and students, making the solution easily replicable and extensible.

Despite these positive aspects, there are several areas in which the proposed system could be improved. First, the training process was relatively long and computationally demanding - successful training required millions of time steps and access to powerful hardware. Furthermore, the neural network operated solely on visual input, which, while reflecting real-world sensory data, increases the difficulty of the learning task.

A combination of visual and structured inputs (e.g., position, orientation, speed) could lead to more efficient learning.

Another potential improvement involves extending the reward function to penalize inefficient driving behaviours (e.g., unnecessary braking or oscillations, driving out of track boundaries), which could further optimize not only the trajectory but also energy efficiency and smoothness.

Lastly, it would be beneficial to experimentally compare PPO with other modern algorithms for continuous control, such as Soft Actor-Critic (SAC) or Twin Delayed Deep Deterministic Policy Gradient (TD3), in order to establish a broader empirical foundation for selecting the optimal strategy.

In accordance with the objective of this project—gaining practical experience in reinforcement learning and preparing for a bachelor’s thesis—the results achieved can be considered successful.

The resulting agent is capable of autonomously and effectively solving a complex simulation task, and the developed infrastructure supports future extensions with additional functionalities or environments.

Conclusion and Implementation

The aim of this study was to design and implement an autonomous agent for the task of vehicle control in the CarRacing-v0 environment using the Proximal Policy Optimization (PPO) reinforcement learning algorithm.

The implementation was carried out in a Windows environment using the Python programming language together with the Stable-Baselines3 and Gymnasium libraries [2], [3]. This work provides a practical demonstration of applying modern RL algorithms in a complex visual simulation setting with a continuous action space.

Training was conducted on a high-performance computing system (Tab.2):

Table 2. Hardware used for training (4 parallel environments).

Parameter	Value
GPU	NVIDIA GeForce RTX 4090 (24 GB VRAM)
CPU	Intel i9-13900K (24 cores, 32 threads
RAM	128 GB DDR4 3600 MHz

The agent was trained in four parallel vectorized environments for a total of 5 000 000 interactions with the environment. The overall training time amounted to 559 minutes and 22 seconds (Tab.1). The resulting model achieved an average score exceeding 500 points per episode, which indicates advanced driving capabilities of the agent. Observations from the testing phase revealed that the agent not only managed to navigate corners smoothly but also demonstrated recovery capabilities in cases of temporary track departure, *actively attempting to reorient and resume driving.*

These results confirm that the combination of PPO and a well-designed training regime enables successful mastery of complex visual and control tasks. The implementation provides a robust foundation for further research, such as extending to more dynamic environments, employing more realistic physics, or integrating multi-agent learning.

Acknowledgement

The author gratefully acknowledges the supportive academic environment that facilitated the development of this project. This work was carried out independently as part of personal research and in preparation for a future bachelor's thesis, with the aim of acquiring practical experience in the field of reinforcement learning.

References

[1] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O. (2017). Proximal Policy Optimization Algorithms. arXiv preprint arXiv:1707.06347.

[2] Raffin, A., Hill, A., Gleave, A., Kanervisto, A., Ernestus, M., Dormann, N. (2021). Stable-Baselines3: Reliable Reinforcement Learning Implementations. Journal of Machine Learning Open Source Software. Retrieved from https://github.com/DLR-RM/stable-baselines3

[3] Farama Foundation (2024). Gymnasium Documentation. Retrieved online, May 2, 2025, from https://gymnasium.farama.org

Authors

Branislav Valacsay
Faculty of Informatics
Pan-European University in Bratislava, Slovakia
xvalacsay@paneurouni.com

Branislav Valacsay is a student of Applied Informatics at the Faculty of Informatics, Paneuropean University in Bratislava. His academic interests include machine learning, artificial intelligence, and reinforcement learning. As part of his independent student research activity, he focused on applying deep reinforcement learning methods to simulation-based control tasks. This work also serves as the foundation for his future bachelor's thesis.

International Journal
Information Technology Applications (ITA)
Volume 14, Number 1, July 2025

OPTIMAL DESIGN OF A HOME NETWORK IN A Wi-Fi ENVIRONMENT

Lukáš Horčičák, Erik Chromý

Abstract:

The primary objective of this study is to design an optimal home wireless network based on an in-depth analysis of Wi-Fi signal coverage and interference. The research focuses on applying IEEE 802.11 standards and practical tools to measure, simulate, and improve the wireless environment in a real multi-floor household. Measurement tools such as NetSpot and iPerf were used to analyze signal strength, data throughput, and noise. Predictive simulations were performed using Hamina software to determine the optimal positioning of access points and frequency planning. The redesigned network was deployed and re-measured to compare performance metrics. Results indicate a significant improvement in signal coverage, more stable connectivity, and increased data transfer rates across all floors. The project also demonstrates the importance of proper channel allocation and network segmentation to reduce interference and improve network security. The presented approach provides a practical methodology that can be replicated in similar home or small office environments.

Keywords:

Wi-Fi, network design, signal analysis, performance optimization, wireless tools.

Introduction

Wi-Fi networks are essential for modern households, yet many suffer from poor coverage and interference due to inadequate design. This article focuses on creating an optimal wireless network using IEEE 802.11 standards and practical tools. A real three-floor house was analyzed with NetSpot for signal strength and iPerf for transmission speed, while Hamina was used for predictive design. The redesigned setup improved performance and stability across all floors.

The network upgrade was planned and executed following the Cisco PPDIOO methodology, which ensured a structured process from initial analysis through implementation and optimization. The results demonstrate that applying a lifecycle-based approach, combined with modern diagnostic and simulation tools, can significantly enhance the reliability and performance of home Wi-Fi networks.

1 The IEEE 802.11 Standard

The IEEE 802.11 standard defines a set of protocols for implementing wireless local area networks (WLANs). Since its introduction in 1997, various amendments have significantly improved throughput, range, and efficiency. These enhancements are primarily achieved through advances in modulation and multiplexing techniques. The earlier standard, 802.11b, used *DSSS (Direct Sequence Spread Spectrum)* with a maximum data rate of 11 Mbps in the 2.4 GHz band.

This is an open access article under the CC BY-NC license
https://creativecommons.org/licenses/by-nc/4.0/

Its successor, 802.11g, introduced *OFDM (Orthogonal Frequency Division Multiplexing)*, raising speeds up to 54 Mbps while maintaining backward compatibility.

802.11n brought *MIMO (Multiple-Input Multiple-Output)* technology, enabling multiple spatial streams for increased capacity. It also supported both 2.4 GHz and 5 GHz bands and introduced *64-QAM* for higher data encoding density.

802.11ac improved upon this by using *256-QAM* and wider channels (up to 160 MHz), offering theoretical speeds over 1 Gbps in the 5 GHz band. The most significant upgrade came with 802.11ax (Wi-Fi 6), which introduced *OFDMA (Orthogonal Frequency Division Multiple Access)*. This method allows multiple devices to share a channel more efficiently, reducing latency and congestion in dense environments. Wi-Fi 6 also added *1024-QAM*, increasing spectral efficiency.

The latest amendment, 802.11be (Wi-Fi 7), further enhances performance with features such as *MLO (Multi-Link Operation)*, 320 MHz-wide channels, and *4096-QAM* modulation. These innovations provide extremely high throughput, ultra-low latency, and improved reliability, particularly for environments with high device density or demanding real-time applications. Though still in early deployment stages, 802.11be is expected to become the new standard for future-proof residential and enterprise wireless networks (Tab.1), [1], [2], [3].

Table 1. Comparison of IEEE 802.11n, 802.11ac, 802.11ax
in terms of speed, frequency bands, technologies used, and typical applications.

Standard	**Max Speed**	**Band(s)**	**Key Techniques**	**Use Case**
802.11n	600Mbps	2.4/5GHz	MIMO, OFDM, 64-QAM	Older networks
802.11ac	~1.3Gbps	5GHz	MIMO, OFDM, 256-QAM	High-throughput media
802.11ax	>9.6Gbps	2.4/5/6GHz	OFDMA, MU-MIMO, 1024-QAM	Modern homes, IoT
802.11be	>40Gbps	2.4/5/6GHz	MLO, 4096-QAM,	Future-ready

Given today's typical household usage, including video streaming, online gaming, smart home devices, and remote work. The recommended standard is IEEE 802.11ax (Wi-Fi 6) or Wi-Fi 6E, which adds support for the 6 GHz band. These standards offer higher throughput, better spectrum efficiency, and are optimized for environments with many connected devices.

Choosing a router that supports Wi-Fi 6 ensures future-proofing, better coverage, and significantly improved performance, particularly in homes with multiple users and smart devices.

2 Wireless Network Modernization Using PPDIOO Methodology

The modernization of a home wireless network was structured according to the Cisco PPDIOO methodology (Prepare, Plan, Design, Implement, Operate, Optimize). This structured lifecycle approach provided a clear framework for each phase of the project.

In the Prepare phase, the technical requirements of the household were analyzed, including the types of devices, their Wi-Fi standards, and typical usage (e.g. streaming, gaming, remote work). The Plan phase included a thorough site survey using tools like NetSpot for passive measurements and iPerf for throughput testing. In the Design phase, the wireless environment was virtually modeled using Hamina software to predict signal propagation and optimize access point placement. The Implementation phase involved the physical deployment of access points and selection of appropriate frequency channels based on spectrum analysis. During the Operate and Optimize phases, the network's performance was monitored and refined, including segmentation for IoT security and improved channel allocation. The PPDIOO framework ensured a systematic and effective network upgrade tailored to real household needs (Tab.1), (Tab.2).

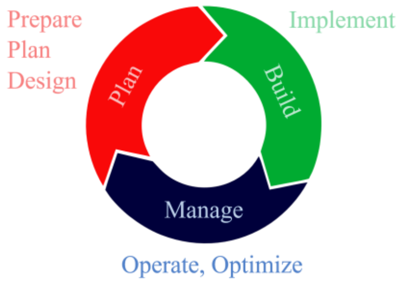


Fig.1: PPDIOO diagram.

2.1 Prepare

In the preparation phase, the wireless needs of the household were identified. The tested environment was a three-story family house equipped with multiple smart devices, including TVs, mobile phones, laptops, and IoT elements such as cameras and smart lighting. Key requirements included full coverage in all rooms, stable performance for 4K streaming, online meetings, and isolated traffic for IoT devices.

2.2 Plan

A detailed survey was conducted using NetSpot for passive signal analysis and iPerf for active hroughput testing. The tested network operated on the IEEE 802.11ac standard, utilizing both 2.4 GHz and 5 GHz frequency bands. Initial RSSI measurements revealed signal drops below –70 dBm in various rooms, especially on the upper floor. The 2.4 GHz band suffered from interference due to overlapping channels, which was confirmed by spectrum analysis (Fig.2-4).

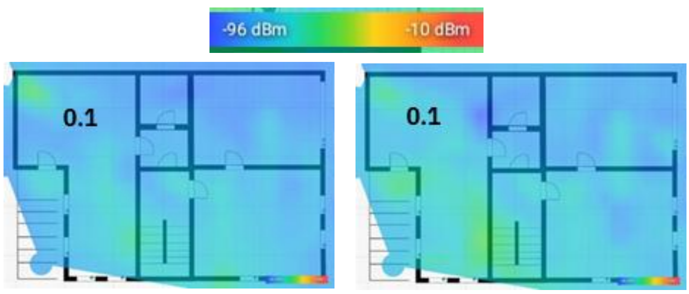


Fig.2: Heatmap - NetSpot RSSI - Basement 2.4 GHz and 5 GHz.

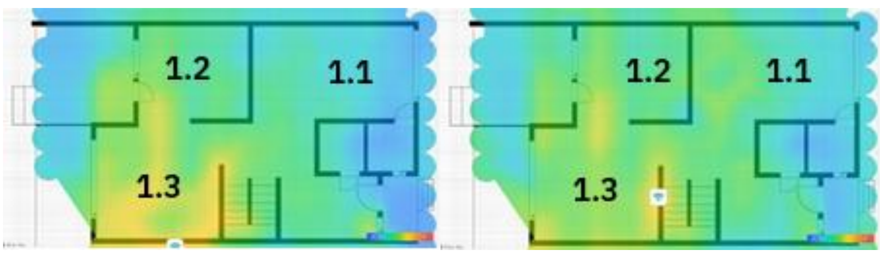


Fig.3: Heatmap - NetSpot RSSI - 1st floor 2.4 GHz and 5 GHz.

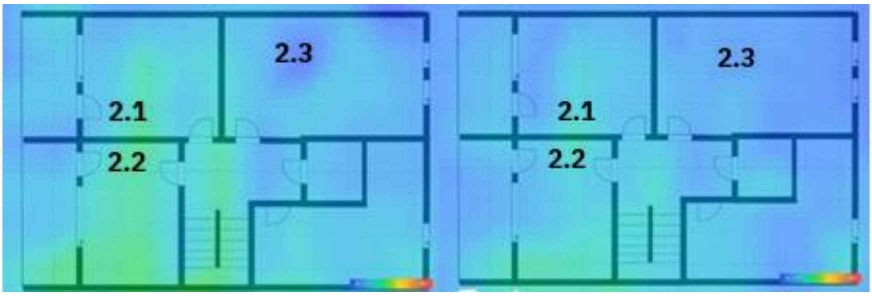


Fig.4: Heatmap - NetSpot RSSI - 2nd floor 2.4 GHz and 5 GHz.

Transmission speed measured using iPerf was:

Table 2: Transmission speed - Before upgrade.

	Frequency band	
Room:	2,4 GHz	5 GHz
1.1	81 Mbit/s	135 Mbit/s
	76 Mbit/s	145 Mbit/s
	94 Mbit/s	145 Mbit/s
1.2	98 Mbit/s	304 Mbit/s
	96 Mbit/s	389 Mbit/s
	97 Mbit/s	349 Mbit/s
1.3	91 Mbit/s	409 Mbit/s
	79 Mbit/s	424 Mbit/s
	92 Mbit/s	477 Mbit/s
2.1	79 Mbit/s	20 Mbit/s
	76 Mbit/s	21 Mbit/s
	90 Mbit/s	13 Mbit/s
2.2	89 Mbit/s	136 Mbit/s
	76 Mbit/s	144 Mbit/s
	94 Mbit/s	122 Mbit/s
2.3	28 Mbit/s	1 Mbit/s
	32 Mbit/s	2 Mbit/s
	31 Mbit/s	1 Mbit/s
0.1	88 Mbit/s	210 Mbit/s
	73 Mbit/s	233 Mbit/s
	87 Mbit/s	192 Mbit/s

2.3 Design

The network was virtually redesigned using Hamina Wireless software (Fig.5-7). Two access points (APs) were strategically placed on each floor to ensure seamless coverage. Non-overlapping channels (1, 6, 11) were selected for the 2.4 GHz band. In the 5 GHz spectrum, DFS channels were used to avoid neighboring network interference.

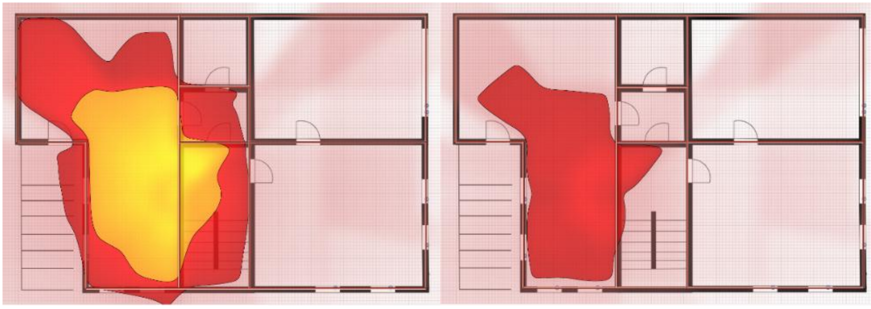


Fig.5: Simulation of Hamina coverage - Basement 2.4 GHz and 5 GHz.

Fig.6: Simulation of Hamina coverage - 1st floor 2.4 GHz and 5 GHz.

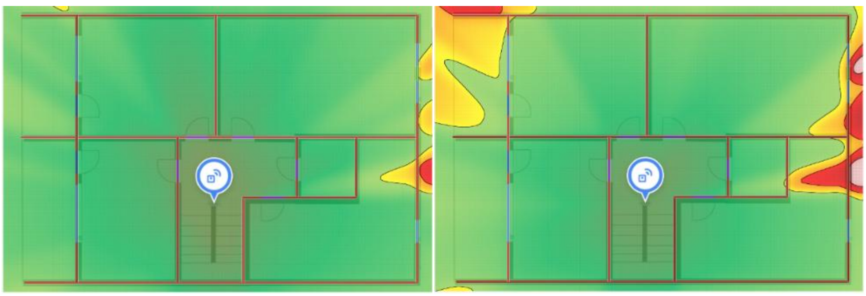


Fig.7: Simulation of Hamina coverage - 2nd floor 2.4 GHz and 5 GHz.

2.4 Implementation

The implementation phase involved physical installation of access points based on the design. Ceiling-mounted APs with Power over Ethernet (PoE) were used to allow flexible placement. Transmission power was adjusted to prevent signal bleeding between floors. Channel allocation and minimum RSSI thresholds were configured to support stable roaming.

2.5 Operate

After deployment, the network was monitored using FortiOS tools to track client distribution, signal strength, and performance. Adjustments included fine-tuning roaming parameters and verifying signal balance across the house.

Tab.3: Transmission speed - After upgrade

	Frequency band	
Room:	2,4 GHz	5 GHz
1.1	116 Mbit/s	426 Mbit/s
	115 Mbit/s	415 Mbit/s
	96 Mbit/s	411 Mbit/s
1.2	138 Mbit/s	865 Mbit/s
	142 Mbit/s	796 Mbit/s
	139 Mbit/s	824 Mbit/s
1.3	182 Mbit/s	932 Mbit/s
	184 Mbit/s	907 Mbit/s
	185 Mbit/s	901 Mbit/s
2.1	90 Mbit/s	753 Mbit/s
	100 Mbit/s	715 Mbit/s
	85 Mbit/s	718 Mbit/s
2.2	89 Mbit/s	909 Mbit/s
	95 Mbit/s	855 Mbit/s
	97 Mbit/s	910 Mbit/s
2.3	57 Mbit/s	492 Mbit/s
	58 Mbit/s	468 Mbit/s
	61 Mbit/s	492 Mbit/s
0.1	80 Mbit/s	209 Mbit/s
	86 Mbit/s	198 Mbit/s
	82 Mbit/s	200 Mbit/s

2.6 Optimize

In the proposed solution, the wireless network was segmented into three separate VLANs to enhance security and manage traffic more efficiently. One segment was dedicated to primary household users, another to guest devices with limited access, and a third isolated VLAN was used exclusively for IoT and smart home appliances. This segmentation helped prevent unauthorized access between device groups and reduced potential interference and network load.

Performance analysis based on iPerf measurements revealed a substantial increase in throughput. As shown in (Fig.9), the most significant improvements were observed in the 5 GHz band, where some locations (e.g., 1.3, 2.3) exceeded 900 Mbps, compared to values below 200 Mbps before optimization. In the 2.4 GHz band, improvements were also evident, with transmission speed nearly doubling in several areas.

This visual comparison confirms the effectiveness of deploying multiple access points with optimized placement and transitioning to the IEEE 802.11ax standard. The redesigned network consistently delivered high throughput and reliable coverage across all tested zones (Fig.8), (Fig.9).

Primary VLAN: Users and trusted devices.
Guest VLAN: Temporary devices
IoT VLAN: Smart home appliances

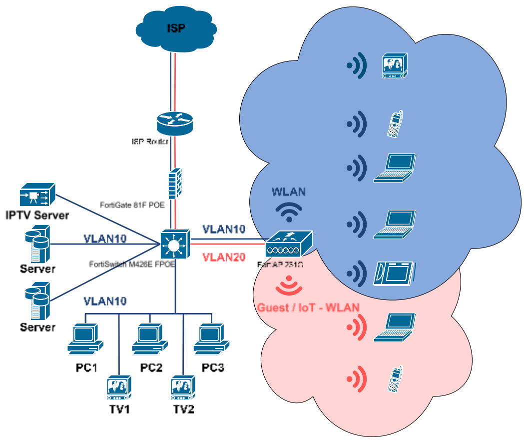


Fig.8: Network segmentation.

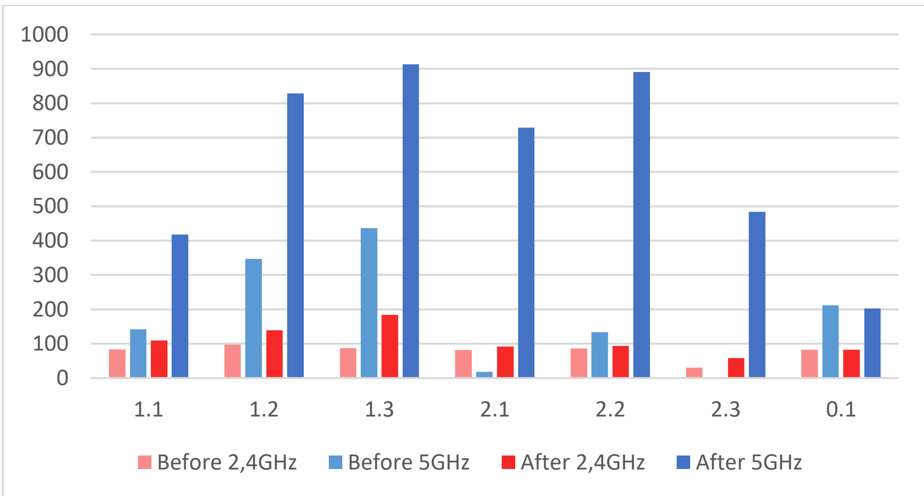


Fig.9: Transmission speed comparison.

Conclusion

This paper presents a practical approach to designing and optimizing a home Wi-Fi network using real-world measurements, predictive simulation, and structured planning based on the PPDIOO methodology. The original network, based on the IEEE 802.11ac standard and a single access point, was upgraded to a multi-access point infrastructure using the IEEE 802.11ax (Wi-Fi 6) standard. This upgrade, combined with optimal placement and network segmentation, significantly improved coverage, throughput, and overall reliability across all floors of the household. The results confirm that applying modern Wi-Fi standards and design tools can greatly enhance wireless performance in residential environments.

Acknowledgement

The author would like to thank to Ing. Erik Chromý, PhD., for his expert guidance and support during the development of this work. Appreciation is also extended to Pan-European University, Faculty of Informatics, for providing the academic environment and resources necessary to complete the project. This work was conducted as part of the bachelor thesis titled Optimal Design of a Home Network in a Wi-Fi Environment.

References

[1] Hucaby, D. (2014). CCNA Wireless 640-722 Official Cert Guide. Cisco Press. ISBN 978-1-58720-561-3.
[2] Edgeworth, B., Henry, J., Ransome, J. (2020). CCNP and CCIE Enterprise Core ENCOR 350-401 Official Cert Guide. Cisco Press. ISBN 978-0-13-678632-0.
[3] Henry, J., Dougherty, D. (2024). CCNP Enterprise Wireless Design ENWLSD 300-425 and Implementation ENWLSI 300-430. Cisco Press. ISBN 978-0-13-745962-9.

Authors

Lukáš Horčičák
Pan-European University, Faculty of Informatics, Bratislava, Slovakia
xhorcicak@paneurouni.com
Lukáš Horčičák is a student of Applied Informatics at the Faculty of Informatics, Paneuropean University in Bratislava.

Ing. Erik Chromý, PhD.
Pan-European University, Faculty of Informatics, Bratislava, Slovakia
erik.chromy@paneurouni.com
Erik Chromý is a specialist in computer networks and telecommunications. At the Faculty of Informatics, he focuses on network quality, modern communication systems, and simulation tools. He has contributed to several applied research projects and regularly publishes in international journals and conferences.

International Journal
Information Technology Applications (ITA)

Volume 14, Number 1, July 2025

CYBERSECURITY CHALLENGES ASSOCIATED WITH SOFTWARE UPDATES

Robert Valík

Abstract:

In the past, software distribution was one of the tougher problems for independent developers. With the rise and ubiquitousness of broadband internet connectivity, this problem has been solved as these developers have the ability to directly distribute software updates to their customers. As the focus is on fast-paced software delivery, they may lack practical knowledge in some other, more subtle aspects of software distribution, e.g. information security.

Our hypothesis is that there is a lack of security auditing before the final distribution. Using a controlled environment, we analyzed outgoing network traffic to identify the communicating during update sequence.

We used acquired data to identify a specific service that serves software update packages. After we identified the service, we tried to force the software to use our modified service, download a modified update package and execute it. As we were able to succeed, this outcome fully supports our above assumptions.

Keywords:

Software updates, security, vulnerability.

Introduction

In the last decades, the software development and distribution process has evolved considerably. In the past, before the abundance of broadband internet connectivity, the software distribution was handled by stone and mortar shops or postal services. One had to buy a physical medium containing the desired software. Delivery of regular patches was on the edge of unthinkable. The focus of the software development was to build and deliver a finalized working product. After the release the work would shift to building a newer version with a new functionality or a different product altogether.

After the emergence of the internet, new pathways for the software delivery became apparent. Potential clients may buy the software online, download it, install and use it nearly instantaneously [1]. If a problem with the software that warrants a new version arises, it is quite easy to download and apply the patch. Most modern applications contain a self-update mechanism that periodically checks if any update is available and install it.

As the updates on the top of new functionalities may contain fixes for newly discovered security issues, prompt installation of the software updates is considered a best practice [2]. But can we consider it to be safe in any situation? We will discuss some possible scenarios where it may have some undesirable effects.

This is an open access article under the CC BY-NC license
https://creativecommons.org/licenses/by-nc/4.0/

1 Hypothesis

The independent software studios or individual software developers have always existed, but the spread of the public / consumer internet access broadened the field for alternative software delivery methods and allowed the independent developers to thrive without the need to worry about the logistics of the distribution of physical installation media.

Our hypothesis is that as independent software developers focus mainly on software development and functional testing, there may be lack of attention to other parts of the software delivery life cycle. Security may be perceived as an unnecessary overhead and their quality assurance processes may have severely limited or entirely missing security auditing. As a result, many unpatched and overlooked security issues may exist. We also suspect that there may be a lack of understanding of the consequences of delivering insecure products on the creator's side.

In this work we will focus on one of the possible attack vectors on the update process.

2 Approach Description

Our intention is to show that some ways to influence a software update process with undesirable outcome for the software user do exist. To manipulate the process, we need to have some insight into it - how it works, where does it connect, which protocols the process uses. There are more ways to achieve this - direct reverse engineering or de-compilation of the software application or black-box testing [3]. We will apply the black-box approach.

The first step will be to passively capture the network traffic, identify relevant protocols, encryption and endpoints for transmission. The result may help us to identify some weak points in communication. There are a lot of tools available already, commercial and open source. We will use the open-source network analyzer - Wireshark.

After successful identification of the communication protocols used, we will try to identify and analyze the update payload, its structure and content.

After the analysis phase, if we gather enough information, we will proceed to the "attack" phase. In this phase we will try to create a modified update package, and we will try to force the application to connect to our specially constructed service endpoint, download the modified package and apply it.

Before we start, we have to select software to analyze. One class of software products that depend on rapid updates are - computer games developed by both large and independent game companies which rely on the recent "early access" paradigm [4]. Releasing a game as "early access" is in the essence - a commercial release of an unfinished product. Potential client pays for the promise that he will get a finished product in the future and while using the unfinished product, and that he may be a part of the development (you pay to be a beta tester of a game, and you can keep your license after the product is finished).

3 Analysis

After careful consideration as the first software to analyze, we have chosen a factory building game developed by an independent game studio. The game has been at the time of analysis under heavy development but quite stable and already available as early access. As we were unable to acquire the studio's consent at this moment, we will not disclose more information about the game and will refer to it as The Game.

3.1 Update Sequence Analysis

As the (Fig.1) shows, the update mechanism first uses DNS to resolve two addresses - updater.thegame.net and dl.thegame.net. After this step it uses a TLS encapsulated communication channel to connect to these two servers. After trying to connect to these addresses, it is apparent that the first is an authentication service and the second is an address of a content delivery network that caches and provides a worldwide accelerated access to the update packages.

The Transport Layer Security (TLS) [5], if implemented correctly, provides robust security over while communicating over other networks. It relies on cryptographic algorithms and if a party provides a cryptographic certificate, the other party can verify the counterpart's identity.

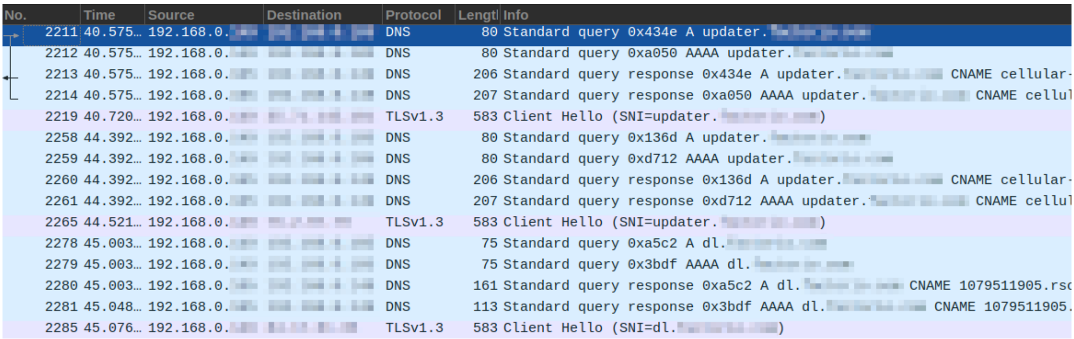


Fig.1. Network traffic captured using Wireshark - DNS queries and TLS connections to the update server are easily visible.

After the package is downloaded, it is applied to the game, the game is restarted automatically, and the player may enjoy it. The whole update flow is visualized on the sequence diagram on (Fig.2).

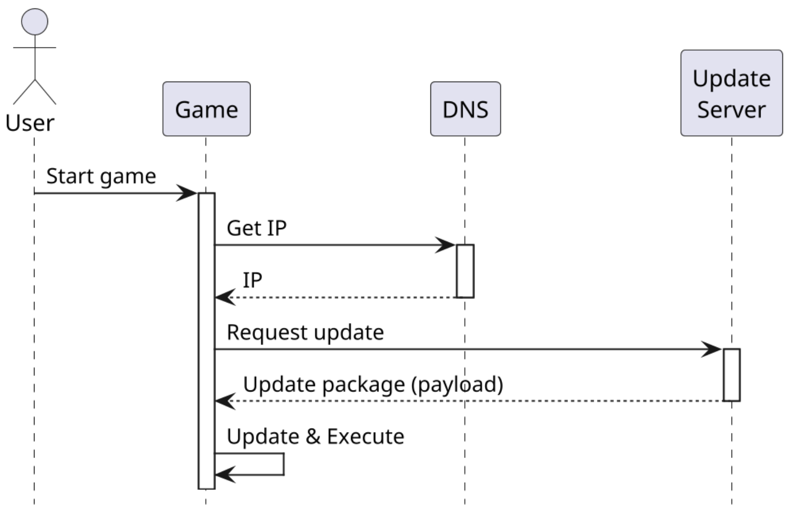


Fig.2. Communication flow during the update process.

To better understand the communication between the game and the update servers, we need to see what is inside of it, and how it works. The easiest way to do it is to prepare a Manipulator-in-the-Middle [6] attack environment to intercept the communication between The Game and the legitimate update server as shown on the (Fig.3).

We will redirect the DNS requests to a DNS server under our control. As the game depends on the operating system to do address resolution, it has no direct authority about it. Our DNS will reply with an IP that redirects the game to our fake update server. This web/proxy server that will transparently forward the requests to their desired destination, logs the whole communication and keeps the transmitted data fragments.

The game successfully connects to our server. The connection to our update server is secured by TLS, but it has no valid certificate as no certificate authority would provide us with a valid certificate without any proof that we are the rightful owners of the domain. (Un)fortunately, the validity of the certificate is never verified - and this is the first security problem, that in principle nullifies the confidentiality of the resulting connection.

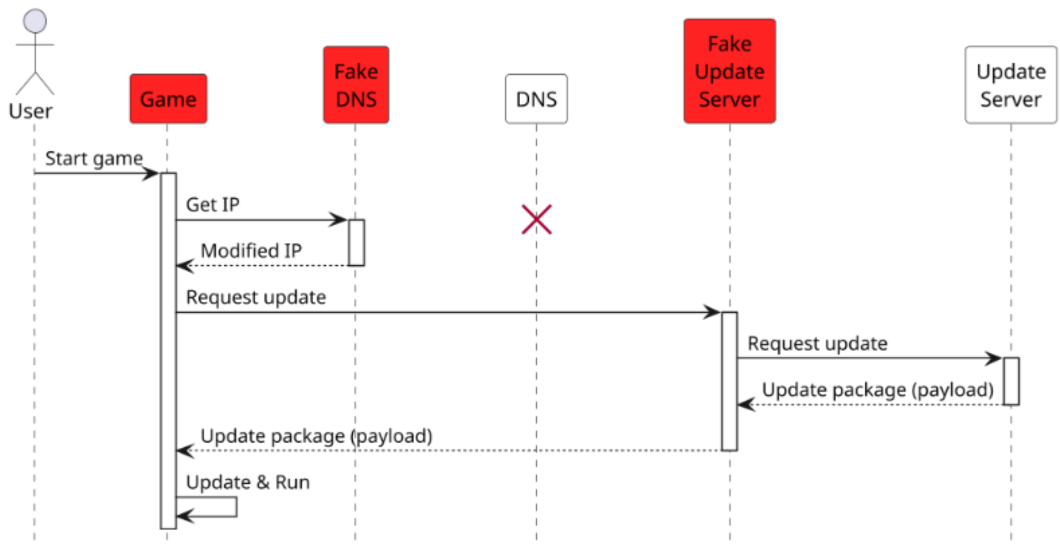


Fig.3. Network traffic during the analysis phase.

As the first step the game checks for available update packages, then downloads an update package for the current version. At this moment we have enough information about the update process and may proceed with the next step - inspection of the update payload.

3.2 Payload Structure

The captured payload is a single ZIP file containing differential data. The ZIP file contains a JSON file describing metadata about the package and all file operations required to update The Game to the new version as shown in (Fig.4). It also contains all the files that need to be updated. One of these files is an executable file thegame.exe that is used to start The Game (Fig.5.).

```
{
  "type": "update",
  "package": "core_expansion-win64",
  "version":
  {
  },
  "apiVersion": 1,
  "membership": "transport-belt-repair-man",
  "author": "[illegible]",
  "contact": "[illegible]",
  "www": "[illegible]",
  "description": "[illegible] update package (core_expansion-win64 2.0.27 → 2.0.28)",
  "files":
  [

    {
      "file": "__PATH__executable__/[illegible]",
      "action": "differs",
      "old_crc": 3885279022,
      "crc": 2396609988
    },

    {
      "file": "__PATH__executable__/[illegible].pdb",
```


Fig.4. Description of the update package in JSON format.

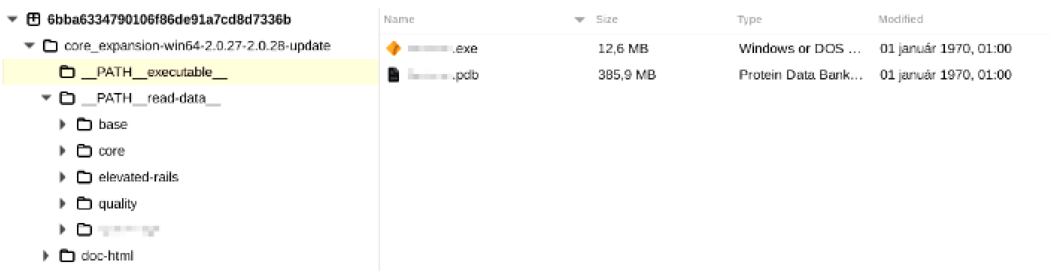


Fig.5. File list of the update ZIP archive.

We expected to find some kind of code signature [7] but we were unable to find it anywhere in the ZIP file or the contained JSON. Only Cyclic Redundancy Check sums are provided for some files. Without this information The Game will be unable to check the validity of the patch. This implies that the game developer has complete trust in the underlying infrastructure or delegates communication security to the often-non-IT-professional, user.

4 Attack Phase

After gathering this information, we may proceed with the “attack” phase. In this phase we will re-use the prepared network environment from the analysis phase. The web server will behave differently and will not resend the update requests but will serve the modified payload instead.

4.1 Payload Modification

We must prepare a modified payload that will pass the internal checks of The Game. Fortunately for us, as we discovered earlier, they are missing. The payload will contain only our specially prepared non-malicious executable file that will display an error message and a simple JSON file describing the installation steps (Fig.6.):

1. delete the existing executable
2. replace the executable with our version


```
  "www": "http://[illegible]",
  "description": "[illegible] update package (core_expansion-win64 2.0.27 → 2.0.28)",
  "files":
  [

    {
      "file": "__PATH__executable__/[illegible].exe",
      "action": "removed"
    },

    {
      "file": "__PATH__executable__/[illegible].exe",
      "action": "added"
    }

  ]
}
```


Fig.6. JSON description of the modified update package.

4.2 Modified Update Process

After the game requests an update, the modified payload is delivered by our server. And this is a second security problem. The implemented update process lacks any authenticity verification, and the payload is trusted without any kind of validation. After the update sequence is completed, the modified payload is executed without any additional user interaction and the target application is compromised. The whole update sequence is shown on (Fig.7).

We have tested this vulnerability on all the major PC operating systems (Windows, macOS and Linux) and succeeded. Even the macOS and Windows built-in security protection against executing an unknown program (Gatekeeper, resp. AppLocker) is not triggered. As the user is aware that an update is in progress, even a privilege escalation may be possible.

4.3 The Fix by the Developer

We have reported the issue to the developer, and it has been fixed a few days later. The update now checks the certificate validity and displays a technical warning that the certificate cannot be verified but older unpatched versions of The Game are still available on the developer's website.

While reporting the issue we stumbled upon an older bug report from 2013 that described the same issue, but the conclusion of the team somehow missed the point as "only authentication token may be disclosed, the bug will be fix in the future", but it finally ended in a "not a bug" basket without any further resolution.

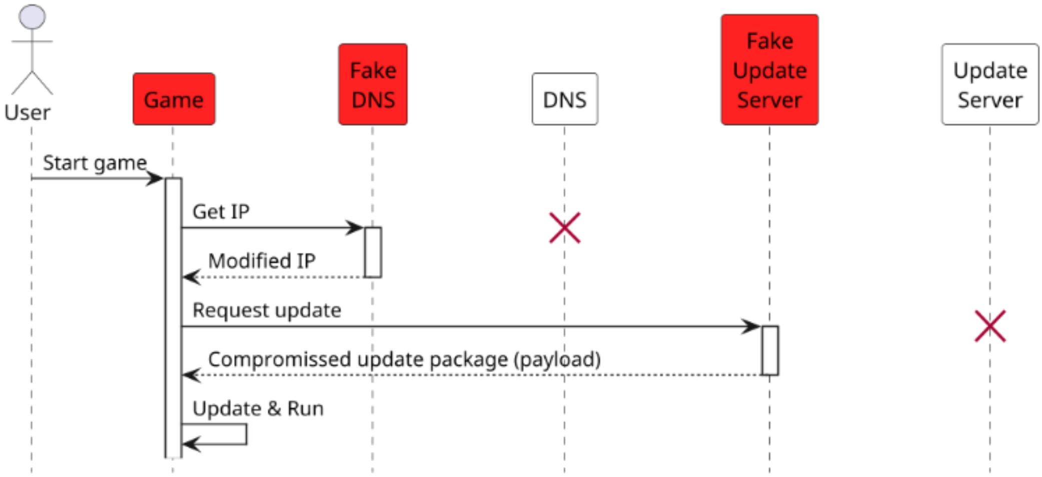


Fig.7. Sequence diagram of the update process with deployment of a modified update package.

Conclusion

The success of this attack confirms our hypothesis about the lack of security focus on the game developer's part, the old bug report even hints that event the consequences of insecure processes are misunderstood. While ignoring the security of the update process, the game may be used as a backdoor to the end-user's computer. As the game is already installed and trusted by the operating system, there is no warning about new executable being installed. If a malicious actor abuses this, the outcome may be quite harsh. User's data may be compromised, deleted, encrypted (ransomware) or a remote access toolkit may be installed.

Information security means balancing the three aspects known as a CIA triad [8] - confidentiality, integrity and availability. Our demonstrations shows what happens if some of these aspects are neglected. To mitigate this type of blunders the game studios should consider employee training in the field of information security and do a regular security audit by a professional third-party auditor.

As this attack relies on the control of the DNS server it is more probable to be deployed in public Wi-Fi scenarios or compromised networks. This implies that any software update should be executed only in a trusted network environment. Also this attack may not be a generic attack but more in the area of spear phishing.

Possible Next Steps

In this work we focused the on a game from the "indie" scene. But what about the A+ titles? How about security updates deployed by different digital distributions or "stores" (Steam, Epic, GOG, Origin, Battle.net, etc.)? And we need to remember that this problem may not be limited to games. This work consisted of a lot of manual steps. Is there a way to automatize it? These are the first areas to consider for further analysis that come to mind, but there may be more.

References

[1] Skelius Mortalis (2024). *Video Game Distribution: From Physical Media to Self-Publishing Breakthrough*. Retrieved online, May 8, 2025, https://www.1d3.com/blog/video-game-distribution-revolution

[2] Gallagher Security Team (2024). *Why software updates are important for security*. Retrieved online, May 8, 2025, https://security.gallagher.com/en/Blog/Why-software-updates-are-important-for-security

[3] Common Attack Pattern Enumeration and Classification (2018). *CAPEC-189: Black Box Reverse Engineering* Retrieved online, May 8, 2025, from https://capec.mitre.org/data/definitions/189.html

[4] Dale Beerling (2020). Alphafunding – *The new trend?* Retrieved online, May 8, 2025, https://www.indiegamemag.com/alphafunding-the-new-trend/

[5] Eric Rescorla, Internet Engineering Task Force (2018). *The Transport Layer Security (TLS) Protocol Version 1.3.* Retrieved online, May 8, 2025, from https://datatracker.ietf.org/doc/html/rfc8446

[6] The Open Worldwide Application Security Project (2025). *Manipulator-in-the-middle attack*. Retrieved online, May 8, 2025, https://owasp.org/www-community/attacks/Manipulator-in-the-middle_attack

[7] Public Key Infrastructure Consortium (2016). *Code Signing Whitepaper - What it Is, Best Practices, and Why its Important* Retrieved online, May 8, 2025, https://pkic.org/uploads/2016/12/CASC-Code-Signing.pdf

[8] Jeroen van der Ham (2021). *Toward a Better Understanding of "Cybersecurity"*. Retrieved online, May 8, 2025, https://dl.acm.org/doi/pdf/10.1145/3442445

Authors

Robert Valík
Pan-European University in Bratislava, Slovakia
xvalik@paneurouni.com

Student of Applied Informatics with over 20 years of experience in Information and Communication Technology in the areas of telecommunications and financial industry.

International Journal
Information Technology Applications (ITA)

Volume 14, Number 1, July 2025

LOCAL STORAGE OF DATA FROM THE PRODUCTION PROCESS

Filip Žemla, Ján Cigánek

Abstract:

This paper presents the design and implementation of a local data storage system for process data acquired from a model of a manufacturing line. The line is controlled by a PLC, for which a custom control program was developed. Process data are collected and visualized using the Promotic SCADA system, which also handles data logging. The application enables both monitoring and basic control of the production line in real time. Collected data are stored in structured text files, organized into folders by time or batch, allowing for easy access and analysis. This work provides an overview of key industrial automation technologies, including PLC programming languages, SCADA systems, and data handling methods. Special focus is given to the Promotic environment and its capabilities in educational and small-scale industrial scenarios. The system was tested on a school-scale production model and proved reliable and suitable for pedagogical use. Challenges related to communication and data integrity were addressed and resolved. The final solution is cost-effective, simple to maintain, and serves as a foundation for further development or integration with external systems.

Keywords:

PLC, SCADA, HMI, Local storage, Process data.

Introduction

In modern industrial environments, the role of process data has evolved from being a mere diagnostic tool to a critical asset for decision-making, traceability, and continuous optimization. As the manufacturing sector embraces Industry 4.0, the integration of intelligent control systems, real-time monitoring, and secure data storage becomes a strategic imperative. One of the key technological enablers of this transformation is the Programmable Logic Controller (PLC)—a robust automation element designed for precise control under harsh industrial conditions. Coupled with SCADA (Supervisory Control and Data Acquisition) systems and HMI (Human-Machine Interface) solutions, PLCs enable the real-time interaction between operators and machines while capturing and distributing valuable process data [1].

Despite the increasing trend toward cloud-based analytics and remote data access, many industrial facilities, especially those in safety-critical or isolated environments, still require localized data storage solutions. These are especially relevant for manufacturing lines that operate autonomously or within secure zones where internet access is limited or deliberately restricted. In such cases, the ability to store, organize, and retrieve process data locally ensures reliability, auditability, and long-term stability without the cybersecurity risks associated with cloud connectivity [1, 6, 8].

This paper presents the design and implementation of a local data storage system for a model production line, developed as part of an academic research project. The solution is based on a laboratory-scale automated assembly line controlled by a Siemens S7-1200 PLC. A Promotic-based SCADA application was developed to visualize and manage operations, and simultaneously log

This is an open access article under the CC BY-NC license
https://creativecommons.org/licenses/by-nc/4.0/

critical process events (e.g., product movements, alarms, user logins) into a structured local file system. The proposed architecture emphasizes modularity, role-based access control, and timestamped data logging, offering a practical template for small-scale industrial deployments and educational purposes. The study aims to demonstrate how localized data storage, when properly designed, can meet both operational and security requirements in modern production systems [8, 9].

1 State of the Art

The efficient management of industrial data requires a synergistic integration of multiple technological layers, including hardware controllers, programming standards, visualization tools, and storage systems. In the context of modern manufacturing, data plays a key role not only in diagnostics and control but also in long-term planning, traceability, and system optimization [1, 2].

Modern production lines are commonly classified based on their degree of automation, which directly impacts complexity, scalability, and labor demands. The three main types are (Fig.1), [2]:

- **Manual assembly lines** - operated entirely by human workers. Materials are handled manually, and each step is performed at a workstation. Typical for low-volume, customized production.
- **Semi-automated lines** - combine human labor with robotic or automated subsystems. Some actions (e.g. transport, machining) are automated, while operators still perform key tasks. Offers a good balance between cost and flexibility.
- **Fully automated lines** - operate with minimal or no human intervention. PLCs or control systems govern all actions. These lines are further divided into:
  - Hard automation: used for fixed-cycle mass production.
  - Flexible automation: adaptable for different products and configurations.

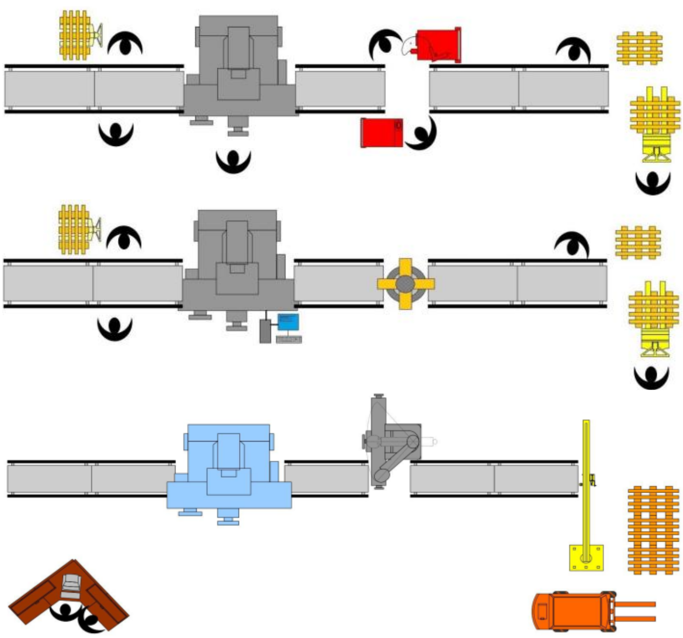


Fig.1. Examples of (up to down): manual, semi-automated, and fully automated assembly lines [2].

At the heart of any automated line lies the **Programmable Logic Controller (PLC)**. PLCs are designed to function reliably in industrial environments, replacing hardwired logic with programmable, reusable control logic. The first PLC, Modicon 084, introduced in 1968, revolutionized industrial automation by enabling digital control of processes [1].

PLCs come in various configurations depending on size and architecture, such as [1]:

- Nano PLCs: < 16 I/O, for simple logic tasks.
- Micro/Mini PLCs: 16–500 I/O, for small production lines.
- Modular or rack-based PLCs: 500+ I/O, used in large-scale industrial systems.

They are also classified by design [1]:

- Compact: fixed number of I/O, integrated in one unit.
- Modular: expandable via I/O and communication modules.
- Rack-mounted: highly scalable, often used in networked systems.

To ensure interoperability and standardized logic, PLC programming is defined by IEC 61131-3, which specifies five primary languages [1]:

- LD (Ladder Diagram) – based on relay logic,
- FBD (Function Block Diagram) – block-oriented,
- IL (Instruction List) – low-level, now deprecated,
- ST (Structured Text) – high-level, similar to Pascal/C,
- SFC (Sequential Function Chart) – process-step driven.

Each one is suited to different control needs. Examples of their syntax and structure are shown in (Fig.2).

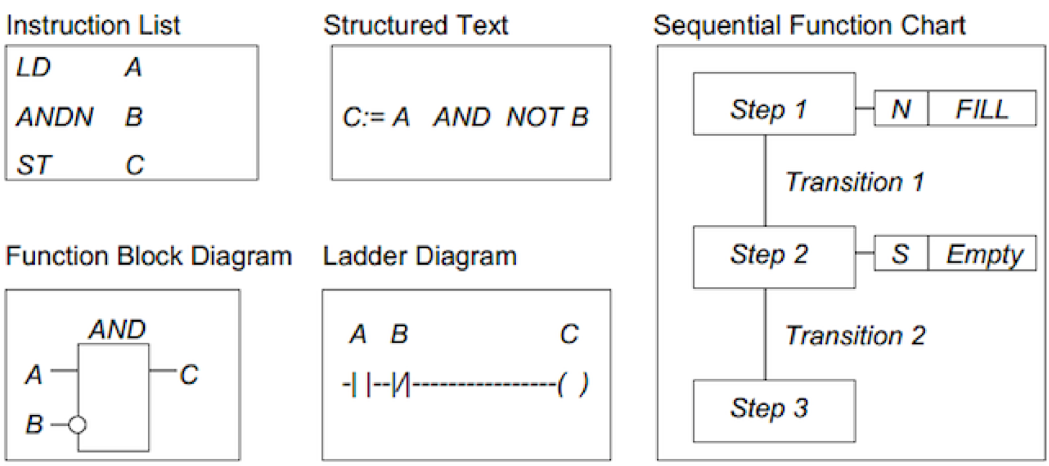


Fig.2. Overview of IEC 61131-3 PLC programming languages [12].

Closely integrated with PLCs are SCADA systems (Supervisory Control and Data Acquisition), which allow real-time visualization and control of automated processes. A SCADA system typically includes [1]:

- Human-Machine Interfaces (HMI).
- Databases or file-based storage.
- Communication interfaces (e.g., OPC-UA, Modbus).

Among commonly used SCADA tools are SIMATIC WinCC, Reliance, and Promotic. In this project, Promotic was selected for its simplicity, licensing flexibility, support for VBScript/JavaScript, and its ability to create and manage local file storage [5].

The Promotic environment transitioned from VBScript to JavaScript in its newer versions. While both are supported, JavaScript is recommended due to broader compatibility. A syntax comparison is provided in (Fig. 3) [5].

VBScript

```
If a > 1 Then
  ' ...
  ' ...
Else
  ' ...
  ' ...
End If
```

JavaScript

```
if (a > 1)
{
  // ...
  // ...
}
else
{
  // ...
  // ...
}
```

Fig.3. Comparison of IF statements in VBScript (left) and JavaScript (right) [5].

From a design perspective, ergonomic principles are essential. The IEC 73 standard defines best practices for SCADA color schemes and operator response. For instance, red indicates critical alarms, yellow means warnings, and green confirms normal conditions. These are summarized in (Tab. 1) [1].

Table 1. IEC 73 color coding for operator interfaces [1].

Color	State	Operator action
Red	Alarm	Immediate reaction required
Yellow	Warning	Monitor closely
Green-yellow	Caution	Attention required
Green	Normal	No action needed

Finally, data storage in SCADA environments is handled either locally or externally. Local storage offers security, autonomy, and stability — especially in air-gapped or isolated systems. External (cloud-based) storage enables remote access and scalability but introduces cybersecurity risks. This work focuses on a local file-based storage architecture, in which process data is written as structured text files into a date-based directory tree. This method ensures transparency, offline functionality and long-term maintainability for small-scale industrial or educational setups [1, 5].

2 Design of Research

The practical implementation was carried out using a laboratory-scale production model designed to emulate the workflow of a real manufacturing line. The aim was to develop a fully functional control, visualization, and data logging system using standard industrial components and software, while ensuring modularity and maintainability.

a) Control Algorithm for PLC

The core of the system is a U-shaped assembly line model by Fischertechnik, comprising four conveyor segments, two processing machines, and five photoelectric sensors for position detection. Transport between line segments is handled by two automated shuttles equipped with motion limit switches. All actuators and sensors are connected through a custom driver interface to a Siemens SIMATIC S7-1200 PLC (CPU 1215C DC/DC/DC), which provides 14 digital inputs and 10 digital outputs, sufficient for the requirements of the experiment (Fig. 4).

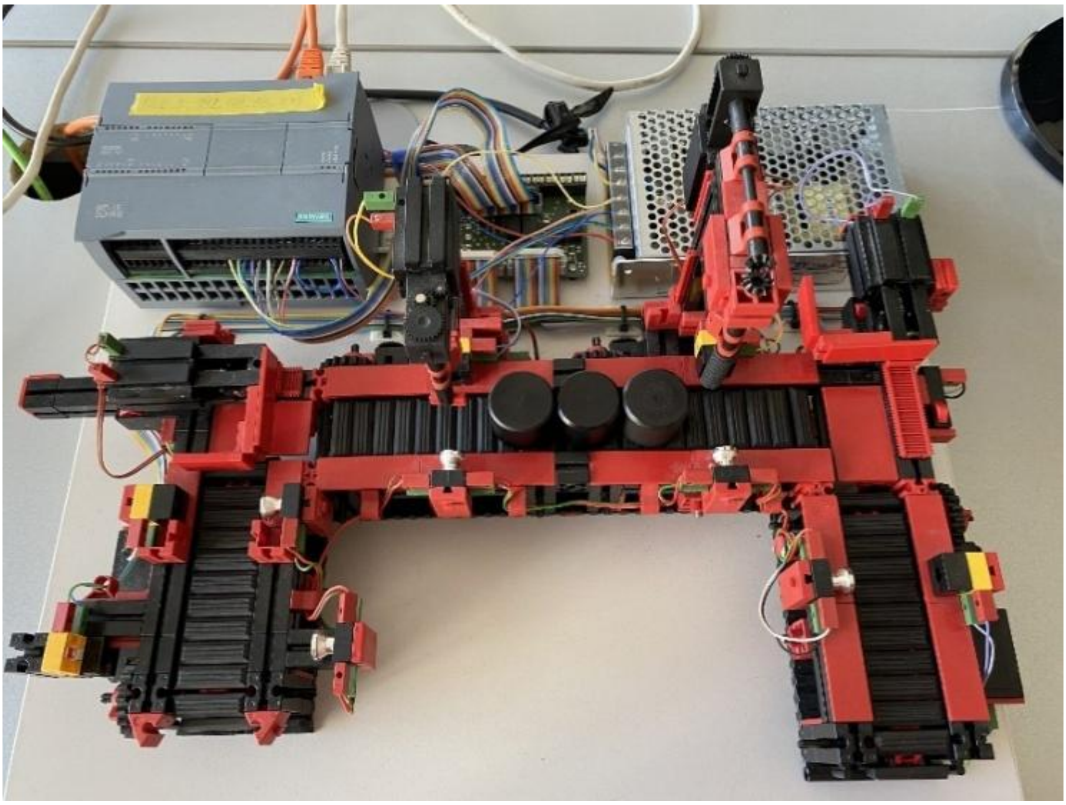


Fig.4. Laboratory U-shaped production line
connected to Siemens S7-1200 PLC via driver interface.

The PLC was programmed using **TIA Portal v17**, utilizing the **FBD (Function Block Diagram)** language due to its intuitive graphical structure and support for real-time logic.

The logic allows two control modes:

- **Automatic mode**, where the PLC executes the entire control logic based on sensor states and time triggers.
- **Manual mode**, which allows authorized users to override machine control via the SCADA interface.

The control logic handles:

- Detection of part positions on the line,
- Operation timing for machine cycles (5s for machine 1, 7s for machine 2),
- Shuttle movement between line sections,
- Counting of finished products and parts currently on the line,
- Triggering and acknowledgment of alarms.

Control mode switching and functional flow of part movement are defined within a hierarchical logic tree (Fig. 5).

b) Human Machine Interface

On top of the PLC logic, a Promotic 8.3.32 SCADA system was developed to visualize and interact with the line. A free license (supporting up to 100 variables and one-hour runtime) was used for development. The SCADA application includes:

- Login panel with user authentication,
- Role-based access levels (monitoring-only, operator, admin),
- Dynamic visualization of line status, material flow, and alarms,
- Manual override panel (admin only),
- Real-time counters and system reset features.

Depending on user level, the main visualization screen (MainPanel) provides different functionalities. The screen layout includes component status (e.g., motor/semaphore color), part counters, alarm state, and manual control buttons (Fig. 6).

Login data, permissions, and authentication logic are handled through an embedded Microsoft Access database. All login attempts, status changes, and alarm confirmations are logged with timestamps and user identifiers. Data generated during the operation of the line are stored in a structured folder hierarchy on the local disk. Each day, a new folder is created and organized by year, month, and date. Within this folder, logs are grouped into two categories:

- Aggregated logs (e.g., system-wide events)
- Product-specific logs (each manufactured part has its own file)

The folder for a given date contains:

- Alarms.txt: a list of all alarm events with timestamps, status, and responsible user.
- Login.txt: a log of all user login, logout, and session change events.
- Material.txt: a summary of finished parts with timestamps of completion.
- ProcessData/: a subfolder where each finished unit receives an individual log file named Material-<ID>.txt.

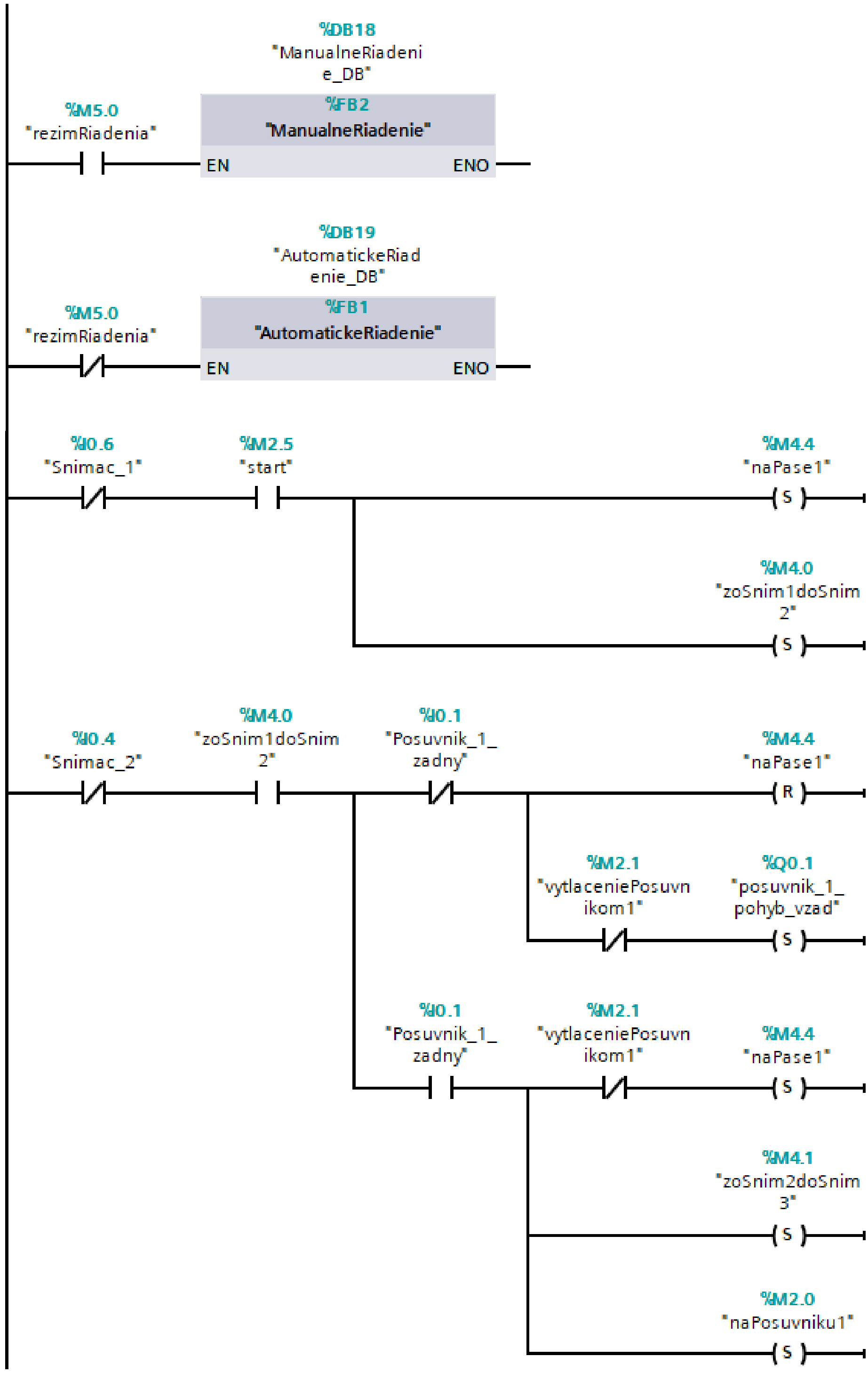


Fig.5. Simplified control logic: mode selection and part handling logic in automatic mode.

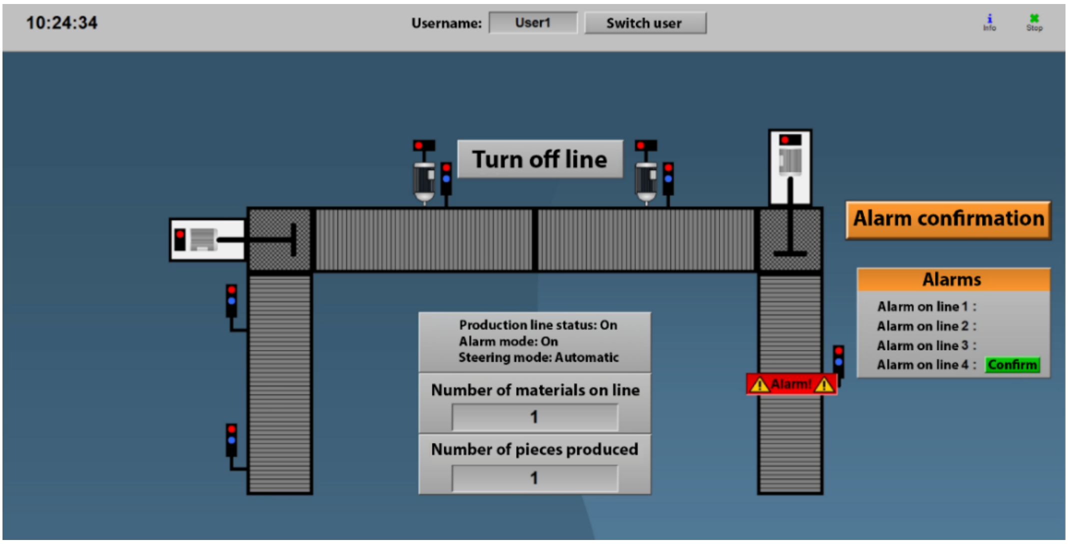


Fig.6. SCADA MainPanel for user level 2 (admin): full system visibility and manual control access.

A sample folder structure is shown in (Fig.7), demonstrating how daily operational data is clearly separated and archived.

```
DatabaseProductionLine
  └── 2024
        └── 5
              └── 2024-05-09
                    ├── Alarms.txt
                    ├── Login.txt
                    ├── Material.txt
                    └── ProcessData
```

Fig.7. Example of folder structure for a specific production day.

The file *Login.txt* captures every authentication-related action, including successful logins, incorrect credentials, user switching, and system shutdowns. All entries are timestamped and labeled by user level (Fig. 8).

Time	- Status	- Name	- Level
05/09/2024 17:01:57	- Login	- User2	- 0
05/09/2024 17:02:03	- Logout		
05/09/2024 17:02:09	- Login	- User1	- 1
05/09/2024 17:02:19	- Logout		
05/09/2024 17:02:26	- Login	- User2	- 0
05/09/2024 17:02:44	- Logout		
05/09/2024 17:02:53	- Login	- User1	- 1
05/09/2024 17:05:59	- Logout		
05/09/2024 17:06:06	- Login	- Admin	- 2
05/09/2024 17:10:24	- Application shutdown and logout		
05/09/2024 17:19:37	- Login	- Admin	- 2

Fig.8. Sample output from *Login.txt* showing login events and user activity.

Alarm handling is also logged continuously. The *Alarms.txt* file records activation, confirmation, and deactivation of each alarm. These logs include alarm name, severity level, status, and the operator who responded. Additionally, *Materials.txt* stores time-marked entries for each completed product (Fig. 9).

Each individual product log contains a full timeline of the part's journey across the production line. This includes:

- Entry and exit timestamps at each sensor.
- Machine start/end cycles.
- Processing durations.
- Final exit from the system.

An example of such a product log is shown in (Fig. 10), illustrating complete traceability from start to finish.

This logging structure provides not only transparency and reproducibility but also makes it possible to perform offline diagnostics and long-term operational analysis without relying on an external database or internet connection.

Time	- Alarm	- Status	- Level	- User
05/09/2024 17:02:38	- Alarm Zone 4	- Active		
05/09/2024 17:02:57	- Alarm Zone 2	- Acknowledged	- 1	- User1
05/09/2024 17:05:22	- Alarm Zone 2	- Active		
05/09/2024 17:05:26	- Alarm Zone 2	- Acknowledged	- 1	- User1
05/09/2024 17:05:36	- Alarm Zone 3	- Active		
05/09/2024 17:05:39	- Alarm Zone 3	- Acknowledged	- 1	- User1
05/09/2024 17:06:16	- Alarms Off			
05/09/2024 17:06:32	- Alarms On			
05/09/2024 17:06:42	- Alarms Off			
05/09/2024 17:06:56	- Alarms On			
05/09/2024 17:07:00	- Alarm Zone 4	- Active		
05/09/2024 17:07:06	- Alarm Zone 4	- Acknowledged	- 2	- Admin

Time	- Number of completed materials
05/09/2024 17:02:57	- Material - 1
05/09/2024 17:05:11	- Material - 2
05/09/2024 17:05:52	- Material - 3
05/09/2024 17:07:06	- Material - 4
05/09/2024 17:07:40	- Material - 5
05/09/2024 17:07:49	- Material - 6
05/09/2024 17:07:59	- Material - 7
05/09/2024 17:08:08	- Material - 8
05/09/2024 17:08:17	- Material - 9
05/09/2024 17:09:00	- Material - 10
05/09/2024 17:20:07	- Material - 11
05/09/2024 17:20:16	- Material - 12
05/09/2024 17:20:25	- Material - 13
05/09/2024 17:20:34	- Material - 14
05/09/2024 17:20:43	- Material - 15

Fig.9. Sample outputs from *Alarms.txt* (top) and *Materials.txt* (bottom).

```
05/09/2024 17:07:17 - Material entered the production process
05/09/2024 17:07:19 - Material arrived at sensor 2
05/09/2024 17:07:29 - Material arrived at sensor 3
05/09/2024 17:07:34 - Machine 1 started
05/09/2024 17:07:37 - Material arrived at sensor 4
05/09/2024 17:07:44 - Machine 2 started
05/09/2024 17:07:49 - Material left the production process
```

Fig.10. Log file for a specific product (*Material-007.txt*) with event history.

3 Discussion of Results and Experiences

The implemented system successfully demonstrated that a combination of industrial-grade control hardware and lightweight SCADA software can be used to design a robust local data storage solution tailored for small production environments or educational settings. Throughout the development and testing process, several key insights were gained regarding both technical functionality and user experience.

From a control standpoint, the use of a Siemens S7-1200 PLC provided sufficient flexibility and reliability for managing a multi-segment production model. The real-time control of conveyors, shuttles, and processing machines worked smoothly in both manual and automatic modes. The system correctly handled concurrent part movements, accurately counted completed units, and responded to unexpected states by triggering alarms. The use of FBD programming proved intuitive and well-suited for this application, enabling efficient troubleshooting and future extensibility.

The Promotic SCADA application played a crucial role in visualization and user interaction. Thanks to its built-in support for scripting, UI design, and data handling, it was possible to develop a fully featured interface without requiring an external database or third-party software. The login system with three user roles ensured appropriate access control and allowed different levels of operational authority, from basic monitoring to full manual override and reset functionality. User feedback indicated that the layout was clear, and the use of color-coded indicators (e.g., semaphores, alarm states) improved usability and situational awareness.

One of the most significant outcomes was the implementation of local file-based data storage, which proved to be both transparent and reliable. The system created structured folders and logs automatically without operator intervention. This approach enabled:

- Traceability of each product (with individual logs for each unit).
- Auditability of system events (e.g., alarms, user logins).
- Offline diagnostics without the need for network access.

Although the use of plain text files may not be suitable for large-scale data analytics or integration with cloud-based ERP systems, it offers several advantages for controlled environments, including air-gapped facilities and student laboratories. Logs could be easily reviewed using basic tools (e.g., text editors or Excel), and the folder structure remained human-readable and well-organized.

During testing, the most challenging aspect was the need to track multiple products simultaneously and assign the correct data to the corresponding product log. This was solved by incrementing unique product IDs based on sensor edge detection and completion counters. The solution worked reliably even in cases where multiple parts were processed in parallel.

Performance-wise, the application was responsive, and the logging mechanism had no observable impact on real-time control, even under frequent data writing intervals (0.1–0.5 s).

The system also responded well to simulated faults and edge cases (e.g., blocked sensors, alarm conditions), confirming the robustness of the control logic and the flexibility of the visualization layer.

In summary, the project confirmed that low-cost SCADA and PLC integration with structured local storage can be used to develop a scalable, secure, and maintainable solution for data collection in automated manufacturing systems. It also offers a useful reference architecture for further expansion into hybrid systems combining local and remote storage or more complex process analytics.

Conclusion

This work presented the design, implementation, and evaluation of a local data storage system for an automated laboratory-scale production line. The project combined industry-relevant technologies—namely: Siemens PLCs, Promotic SCADA software and a custom file-based logging architecture—to create a fully autonomous and offline-capable system for control, monitoring and data archiving.

The solution was built to reflect real-world industrial workflows while maintaining modularity and transparency. The control logic allowed seamless switching between automatic and manual operation, supported alarm handling, and enabled real-time tracking of material flow across the line. The visualization layer, developed in Promotic, provided a clear, role-based interface with differentiated access for operators and administrators.

One of the most important contributions of this project is the implementation of structured local storage that requires no internet connection, external database, or third-party services. All relevant process data—product movements, system alarms, and user actions—are logged in human-readable files, organized by date and type. This not only ensures long-term maintainability but also addresses key cybersecurity concerns by keeping the system fully offline and self-contained.

The solution was tested extensively and proved to be reliable, scalable within its scope, and user-friendly. It offers a viable reference model for:

- Educational labs simulating real production environments.
- Small-scale industrial operations requiring traceability without cloud dependencies.
- Isolated or secure facilities where external data transmission is undesirable.

All project objectives were met, and the system behaved as expected under both normal and fault conditions. Furthermore, the experience gained during development suggests several promising directions for future research, including:

- Integration with external databases or web-based dashboards.
- Migration from flat-file storage to structured formats (e.g., JSON, SQLite).
- Implementation of remote diagnostics and predictive analytics.

In conclusion, the project demonstrates that it is feasible to build a compact and secure data acquisition platform using readily available automation tools and simple design principles. Such systems can contribute to increased transparency, process understanding, and educational value in the field of modern industrial automation.

Acknowledgement

This research was funded by the Slovak Agency for Research and Development - grant No. APVV-21-0125, by the Cultural and Educational Grant Agency of the Ministry of Education, Research, Development and Youth of the Slovak Republic KEGA 021STU-4/2024, and by the Scientific Grant Agency of the Ministry of Education, Research and Sport of the Slovak Republic No. 1/0637/23 and 1/0045/25.

References

[1] Cigánek, J., Žemla, F.: PLC systémy v mechatronike. (*PLC systems in mechatronics*) 2023. ISBN 978-80-8208-097-4.

[2] Kuždák, V.: Optimalizácia výrobnej linky ako súčasť výrobného procesu. (*Production line optimization as part of the production process*) [online]. https://www.sjf.tuke.sk/umpadi/taipvpp/2008/index.files/Priemyselne_inzinierstvo/kuzdak.pdf

[3] Duchoň, F. Babinec, A. Dekanová, M. Mudrák, M. Korec, S.: Metodológia na testovanie nábehu výrobných liniek (1). (*Methodology for testing the start-up of production lines*) [online]. https://www.atpjournal.sk/novetrendy/metodologia-na-testovanie-nabehu-vyrobnych-liniek-1.html?page_id=27597

[4] BPS - Bratislavský parkovací systém. Parkovanie v Bratislave. (*Bratislava parking system. Parking in Bratislava*) [online], January 15, 2025, https://paas.sk

[5] Promotic. Promotic [online]. https://www.promotic.eu/cz/

[6] Hegedűs, T., Varga, I. and Abonyi, J.: Development of a SCADA system using open-source technologies. Periodica Polytechnica Electrical Engineering and Computer Science, vol. 63, no. 1, pp. 14–21, 2019. DOI: 10.3311/PPee.13279

[7] Leitão, P., Karnouskos, S., Ribeiro, L., Lee, J., Strasser, T. and Colombo, A.W.: Smart agents in industrial cyber–physical systems. Proceedings of the IEEE, vol. 104, no. 5, pp. 1086–1101, 2016. DOI: 10.1109/JPROC.2016.2521931

[8] Wollschlaeger, M., Sauter, T. and Jasperneite, J.: The Future of Industrial Communication: Automation Networks in the Era of the Internet of Things and Industry 4.0. IEEE Industrial Electronics Magazine, vol. 11, no. 1, pp. 17–27, 2017. DOI: 10.1109/MIE.2017.2649104

[9] El-Halwagi, M.E.: Design of a SCADA System for Real-Time Industrial Process Monitoring and Control. International Journal of Computer Applications, vol. 182, no. 4, pp. 21–26, 2018.

[10] Swapan, B.: Plant intelligent automation and digital transformation. 2022. ISBN 978-0-323-90246-5.

[11] Gérer, A. et al.: Montážna linka sa spája s menom Ford. (*The assembly line is associated with the Ford name*) In: AT&P journal. ISSN 1336-233X, 2003, vol. 9, p. 12-13.

[12] Motion Control Tips a Design World Resource What Are IEC 61131-3 and PLCopen [online]. https://www.motioncontroltips.com/iec-61131-3-plcopen/

Authors

Ing. Filip Žemla, PhD.
Faculty of Electrical Engineering and Information Technology,
Slovak University of Technology in Bratislava, Slovakia
filip.zemla@stuba.sk
Researcher and full-stack developer. He received the diploma and PhD. degree in Automatic Control from the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology (FEI STU) in 2023. His current work focuses on the virtualization and optimization of modern manufacturing processes, with a strong background in SCADA systems, database technologies, back-end development and front-end development. In addition to his academic background, he actively works as a full-stack developer, combining industrial automation with advanced software engineering skills.

Ing. Ján Cigánek, PhD.
Faculty of Electrical Engineering and Information Technology,
Slovak University of Technology in Bratislava,Slovakia
jan.ciganek@stuba.sk
He was born in 1981 in Malacky, Slovakia. He received the diploma and PhD. degree in Automatic Control from the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology (FEI STU) in Bratislava, in 2005 and 2010, respectively. He is now Assistant Professor at Institute of Automotive Mechatronics FEI STU in Bratislava. His research interests include optimization, robust control design, computational tools, SCADA systems, big data, and hybrid systems.

REPORT OF 32nd INTERNATIONAL CONFERENCE 2025 CYBERNETICS & INFORMATICS (K&I)

Mikulov, Czechia, February 2-5, 2025

organised by
Slovak Society of Cybernetics and Informatics

under the auspices of
Faculty of Electrical Engineering and Information Technology,
Slovak University of Technology in Bratislava, Slovakia

The conference focuses on presentation of latest development in control engineering, information technologies and related multidisciplinary fields in line with current development trends in Industrial Internet of Things (IoT). The conference provides a general forum for researchers, university teachers and users dealing with practice of control and Information Communication Technologies (ICT).

Conference papers may range from theoretical works to engineering applications. This conference is held every two years, changing various places in Slovakia and Czechia countries. The next event should take place in 2027.

Papers are invited within the following fields:

K&I '25 Mikulov

- Methods and algorithms for modelling and control.
- New information and communication systems.
- Embedded, distributed and networked control systems.
- Cyber-physical systems, cloud computing, Big Data and extended reality.
- Artificial Intelligence.

The IEEE conference 2025 CYBERNETICS & INFORMATICS (K&I'25) held in Mikulov, Czechia from January 2^{nd} to January 5^{th} 2025 is at the same time the 32^{nd} International conference CYBERNETICS AND INFORMATICS traditionally organized since 1967 by the Slovak Society for Cybernetics and Informatics at the Slovak Academy of Sciences (SSKI), the national representative of Slovakia in the International Federation of Automatic Control (IFAC) and the European Association for Artificial Intelligence (EurAI). Conference webpage and proceedings:

https://ki2025.sski.sk/
https://ieeexplore.ieee.org/xpl/conhome/10915751/proceeding

ISBN 979-8-3315-4180-4
IEEE catalog number CFP25E30-USB

The K&I'25 conference is organized in cooperation with the Faculty of Electrical Engineering and Information Technology of the Slovak University of Technology in Bratislava, Institute of Informatics at the Slovak Academy of Sciences and supported by important Slovak companies engaged in the research, development and application of complex control systems, informatics and communications, IoT and artificial intelligence.

The conference focuses on presentation of advanced control methods and algorithms, their integration into modern plants with the support of information and communication systems, soft computing and virtual and augmented reality techniques. The aim is to provide a general forum for researchers, university teachers and users dealing with practical problems of control and ICT.

The K&I'25 conference attendees were researchers and teachers from Bulgaria, Czech Republic, Slovak Republic, South Korea and Quatar.

We believe that the IEEE conference 2025 CYBERNETICS & INFORMATICS (K&I'25) significantly contributed to the exploitation and dissemination of research and development results in automation, cybernetics, mechatronics, informatics and communication technologies, to increase the professional level of its attendees and support their scientific work.

We would like to thank all attendees for their inspiring contributions, participation in the conference and fruitful professional as well as informal discussions.

Invited Talks of the K&I Conference

Petri Nets and Object-Centric Processes in Software Engineering

Gabriel Juhás
Faculty of Informatics, Pan-European University, Bratislava, Slovakia

Abstract: In this plenary talk, we discuss the role of Petri Nets and Object-Centric Processes in recent development of Software Engineering. Namely, we illustrate how extended Petri nets can be used to define a low-code language for modelling and development of software systems and web applications based on object-centric processes. Among others, we define key requirements for a low-code language for object-centric processes with forms. Such a language should enable to model object-centric processes consiting of data attributes, the life cycle of data attributes given by a workflow of tasks, and forms linked to these tasks consisting of subsets of data attributes. In addition, we specify requirements for handling events related to artifacts of object-centric processes, namely process instances, data fields, and tasks, resulting in a class of discrete event systems. We also discuss how roles specifying which users can trigger events should be represented within such a language. Finally, we outline requirements for a query language capable of searching on object-centric processes and their tasks. We also provide a prototype of such a language based on Petri nets enriched by data attributes, called Petriflow. Finally, we give an illustrative example of an application implemented using presented low-code language Petriflow.

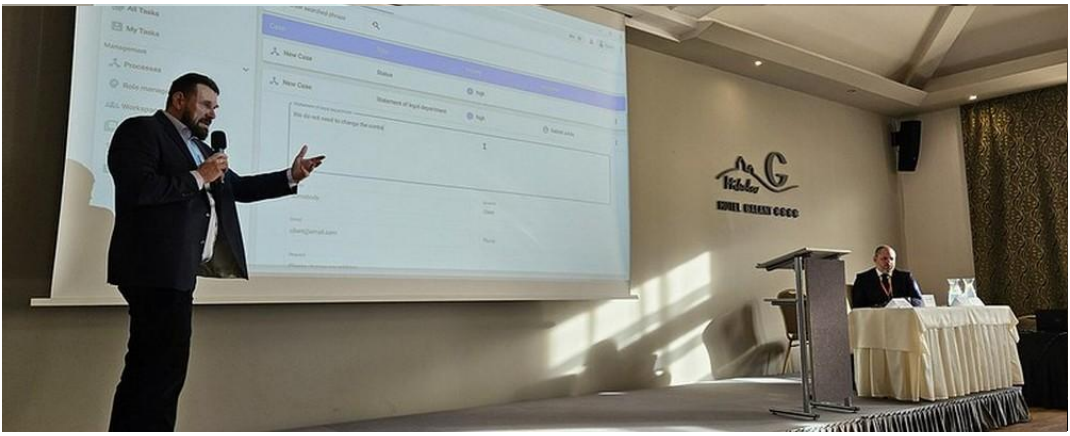
Adapting Technology to the Post-Quantum and Hybrid Warfare Era: Key Research Trends and the Role of the 6G PHYSEC Project

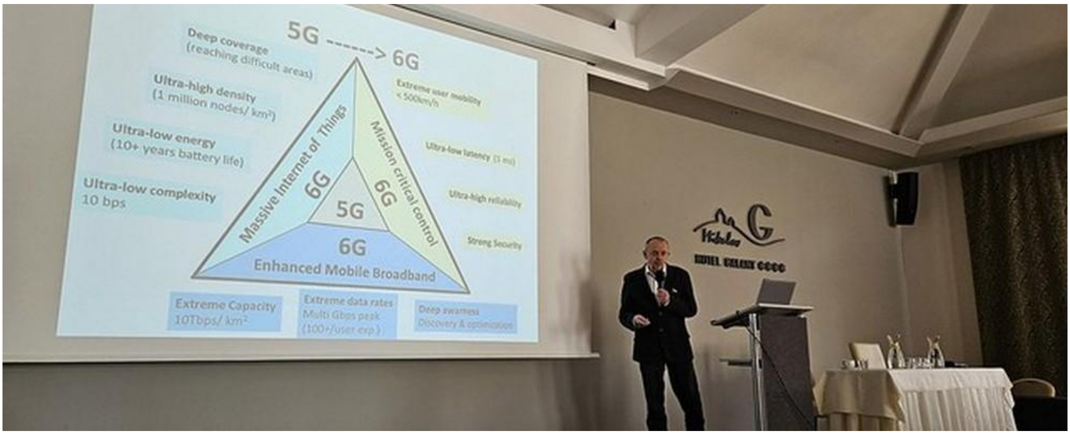
Peter Farkaš
Institute of Multimedia Information and Communication Technologies,
Faculty of Electrical Engineering and Information Technology,
Slovak University of Technology in Bratislava, Slovakia

Abstract: This plenary talk addresses the primary challenges of adapting digital technologies to the post-quantum era in connection with development of 6G infrastructures. Particularly it highlights the research problems in the area of 6G massive machine-type communication security, as well as the protection of information in cloud environments. The talk outlines current research trends that aim to mitigate these challenges and pave the way for the successful implementation of 6G technologies. These include:

- Regeneration coding for cloud systems.
- Physical layer security.
- Integrated sensing and communication.
- Quantum key distribution.
- Quantum communication.

Finally, the objectives of the COST Action CA22138 project, titled Physical Layer Security for Trustworthy and Resilient 6G Systems (6G PHYSEC), will be presented. The project aims to address several of the aforementioned challenges and contribute to the development of secure and resilient 6G systems.




5G ------> 6G
Deep coverage
(reaching difficult areas)
Ultra-high density
(1 million nodes/ km²)
Ultra-low energy
(10+ years battery life)
Ultra-low complexity
10 bps
Extreme user mobility
< 500km/h
Ultra-low latency (1 ms)
Ultra-high reliability
Strong Security
Massive Internet of Things
Mission critical control
6G
6G
5G
6G
Enhanced Mobile Broadband
Extreme Capacity
10Tbps/ km²
Extreme data rates
Multi Gbps peak
(100+/user exp.)
Deep awareness
Discovery & optimization

List of Reviewers

Issue 1/2025, in alphabetic order

Ing. Ivana Budinská, PhD.	Slovak Academy of Sciences, Slovakia
doc. Ing. Oto Haffner, PhD.	Slovak University of Technology in Bratislava, Slovakia
Ing. Michal Hlavatý	Slovak University of Technology in Bratislava, Slovakia
Ing. Zuzana Janatová, PhD.	Slovak University of Technology in Bratislava, Slovakia
doc. Ing. Tomáš Páleník, PhD.	Slovak University of Technology in Bratislava, Slovakia

ISSN: 2453-7497 (online)
ISSN 1338-6468