COMPUTER SCIENCE AND INTERDISCIPLINARY RESEARCH JOURNAL

Volume 2, Issue 2, 2025
ISSN: 3033-1218

https://csirjournal.com

Computer Science and Interdisciplinary Research Journal [CSIR Journal]

ISSN: 3033-1218
Volume 2, Issue 2, 2025.

Published by Science, Education, Technologies Association
Plovdiv, Bulgaria

Licensed under CC BY-NC-ND 4.0.

The views expressed in the articles are those of the authors and do not necessarily reflect the views of the publisher or editorial team.

Editorial Team

Editors-in-Chief

Section COMPUTER SCIENCE

Prof. Asya Stoyanova-Doycheva, Ph.D.

University of Plovdiv Paisii Hilendarski, Plovdiv, Bulgaria

Section MATHEMATICS

Assoc. Prof. Mariyan Milev, Ph.D.

Sofia University St. Kliment Ohridski, Sofia, Bulgaria

Section EDUCATION

Prof. Todorka Glushkova, Ph.D.

University of Plovdiv Paisii Hilendarski, Plovdiv, Bulgaria

Section INTERDISCIPLINARY RESEARCH

Prof. Stanka Hadzhikoleva, Ph.D.

University of Plovdiv Paisii Hilendarski, Plovdiv, Bulgaria

Editorial Board

Prof. Aldo Tagliani, Ph.D.

Università di Trento, Trento, Italy

Prof. Igor Britchenko, Ph.D.

University of the National Education Commission, Krakow, Poland

Prof. Boyan Zlatanov, D.Sc.

University of Plovdiv Paisii Hilendarski, Plovdiv, Bulgaria

Prof. Emil Hadzhikolev, Ph.D.

University of Plovdiv Paisii Hilendarski, Plovdiv, Bulgaria

Prof. Desislava Paneva-Marinova, Ph.D.

Bulgarian Academy of Sciences, Sofia, Bulgaria

Prof. Miglena Koleva, Ph.D.

University of Rousse, Ruse, Bulgaria

Prof. Dobrina Jandik, Ph.D.

University of Arkansas, Fayetteville, United States

Prof. Velika Kuneva, Ph.D.

Agricultural University, Plovdiv, Bulgaria

Assoc. Prof. Senthil Kumar, Ph.D.
Mepco Schlenk Engineering College, Sivakasi, India

Assoc. Prof. Margarita Terziyska, Ph.D.
University of Food Technology, Plovdiv, Bulgaria

Assoc. Prof. Stoyan Cheresharov, Ph.D.
University of Plovdiv Paisii Hilendarski, Plovdiv, Bulgaria

Assoc. Prof. Nikola Shakev, Ph.D.
Technical University of Sofia, Plovdiv Branch, Plovdiv, Bulgaria

Assoc. Prof. Silvia Gaftandzhieva, Ph.D.
University of Plovdiv Paisii Hilendarski, Plovdiv, Bulgaria

Assoc. Prof. Hristina Kulina, Ph.D.
University of Plovdiv Paisii Hilendarski, Plovdiv, Bulgaria

Assoc. Prof. Mikhail Kolev, Ph.D.
University of Architecture, Civil Engineering and Geodesy, Sofia, Bulgaria

Assoc. Prof. Hristo Hristov, Ph.D.
University of Plovdiv Paisii Hilendarski, Plovdiv, Bulgaria

Assist. Prof. Shih-Hao Chang, Ph.D.
National Taipei University of Technology, Taiwan

Assist. Prof. Gordana Rakić, Ph.D.
University of Novi Sad, Serbia

Assist. Prof. Kostadin Yotov, Ph.D.
University of Plovdiv Paisii Hilendarski, Plovdiv, Bulgaria

English Proofreader

Assoc. Prof. Vanya Ivanova, Ph.D.
University of Plovdiv Paisii Hilendarski, Plovdiv, Bulgaria

Design Assistance

Assist. Prof. Maya Stoeva, Ph.D.
University of Plovdiv Paisii Hilendarski, Plovdiv, Bulgaria

Technical Assistance

Ph.D. student Maria Gorgorova
University of Plovdiv Paisii Hilendarski, Plovdiv, Bulgaria

Table of contents

Model for Building a PropTech Ecosystem Through a Network of e-Platforms and Software Service Integration 2

Simeon Kondov, Sylva Pavlova, Hristo Hristov

Cloud-Based Educational Transformation: The Municipal Policy Model of Plovdiv, Bulgaria 10

Stefan Stoyanov

Integrating Economic Indicators and Machine Learning for Exchange Rate Forecasting 18

Ivaylo Boyoukliev, Vanya Ivanova, Hristina Kulina, Boyan Zlatanov

Integration of GenAI and Behavioral Analysis in an Authentication System 26

Georgi Atanasov, Kostadin Yotov, Emil Hadzhikolev

Computer Science and Interdisciplinary Research Journal
Volume 2, Issue 2, 2025. ISSN: 3033-1218, DOI: 10.70862/CSIR.2025.0202-14

Model for Building a PropTech Ecosystem Through a Network of e-Platforms and Software Service Integration

Simeon Kondov [1], Sylva Pavlova [2], Hristo Hristov [3]

[1, 2] PropTech Bulgaria, 26, Industrialna Str., Radinovo Indistrial Zone, 4202 Plovdiv, Bulgaria
[3] University of Plovdiv Paisii Hilendarski, 236, Bulgaria Blvd, 4027 Plovdiv, Bulgaria

[1] simeonkondov2@gmail.com, [2] sylviapavlova@gmail.com, [3] hth@uni-plovdiv.bg

Abstract – The present article examines the conceptualization of a model for building a digital business ecosystem through design and implementation of an e-platforms' network that integrates software services within PropTech Bulgaria as an international organization. "Business Model Archetypes" as a conceptual framework is being applied to the real business case of PropTech Bulgaria, thus clearly showcasing how different business models (BM) correspond to respective e-platforms starting from the basic BMs like product, service, and commerce; developing into combined BM forms such as subscription, marketplace, and brokerage. The system architecture is implemented through a modular approach, led by the understanding that strategies, knowledge, and technologies change over time. As a natural consequence, transaction points change as well. Having a modular structure presupposes a bigger number of transaction points. Currently, there is no communication among separate modules through data transfer. The study advances a practically applicable model for developing a resilient international PropTech ecosystem, i.e., the most complex business model archetype.

Keywords – e-platform, business model, proptech, digital ecosystem, software architecture.

DOI: https://doi.org/10.70862/CSIR.2025.0202-14

Corresponding author: Hristo Hristov,
University of Plovdiv Paisii Hilendarski
Email: hth@uni-plovdiv.bg

Received: 26 June 2025.
Revised: 25 August 2025.
Accepted: 9 September 2025.
Published: 9 September 2025.

© 2025 Simeon Kondov, Sylva Pavlova, Hristo Hristov. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

The article is published with Open Access at http://csirjournal.com/

1. Introduction

In the last decade, software platforms have been established as a strategic instrument for elaborating scalable business models. Progressed by the ever expanding need to connect services, users, and partners globally, e-platforms assume an increasingly central role in structuring business ecosystems. The present article addresses the issue of applying an ecosystem building model via the design and implementation on a modular principle of an e-platforms' network through integrating software services at PropTech Bulgaria. The core business activities of PropTech Bulgaria feature the international business development of PropTech innovations, i.e., innovations in the Construction industry and the Real Estate sector, inclusive of Smart Cities, Smart Buildings, Smart Homes, urban planning innovations, and ESG technological solutions required by the EU Taxonomy legislation enforced in 2021. The study focuses on building an operational international PropTech ecosystem driven by a modular architectural approach, multi-sided business models, and technologically mediated interactions. Both the theoretical foundations of e-platforms and their practical manifestations in the activities of PropTech Bulgaria are explored.

2. Theoretical Context and Relevance of the Problem

The development of e-platforms as a primary software expression of business ecosystems is interlinked with a radical change in comprehending the value of technologies and their application. In the contemporary digital economy, e-platforms such as Amazon, Airbnb, Alibaba, and the alike not only

support commerce or services but design and set the stage for a holistic environment to entertain interaction among multiple stakeholders. All these platforms expose an exponential effect being "multisided platforms," as defined by Zhao et al. [1]. Junic Kim [2] describes such platforms as economic platforms, or intermediary platforms, as these are characterized by multiple individual players/stakeholders who mutually benefit from staying in a common network. Kim [2] underscores one peculiarity of such platforms, namely, the platform owner incurs costs to serve multiple participant groups and collects revenue from various stakeholders involved in the platform. In this particular aspect, Eisenmann et al. [3] further explain that the platform business model fundamentally differs from the linear business model of conventional businesses. In the latter case a company pays suppliers for raw materials and receives money from customers for a finished product/service; customers and suppliers are clearly defined and identifiable. However, this is not the case with platforms: different participants can simultaneously be suppliers and customers to other participants, and the platform owner merely facilitates, maintains, and regulates the network to enable more transactions.

New economic dependencies are being formed in such an environment, the exchange of value is accelerated, and resilient networks of intertwined services are being constructed. The concept of platforms as "an interface of a business ecosystem" [4] implies a new paradigm which formulates the platform not limited to just a technical product, but extended to a means intended to structure the interaction among participants within a specific market or industry. As Cusumano [5] concludes, "Who wins and who loses these competitions is not simply a matter of who has the best technology or the first product. It is often who has the best platform strategy and the best ecosystem to back it up." We can hitherto highlight that the presence of a business ecosystem backing up a platform is integral to the success of the platform. According to Moore [6], there is yet another dimension that impacts the platform strategy success, namely the life-cycle stage the respective company is in, duly reflected in its value proposition. He accentuates that the functionalities (i.e., the type) of a platform change in accordance with life-cycle stage. Therefore, the Business Model Archetypes tool as applied to PropTech Bulgaria's case, outlines in Figure 1 the trajectory of this movement/development of the company's web platforms. Further, Table 1 below projects this movement over the years.

Being the author of the business ecosystem concept, Moore [6] defines it in 1993 as a loosely connected business community incorporating various levels of organizations (industry players, associations, government organizations, etc.), which share a common goal, and aim at achieving mutual development in a strive to cope with the fluctuating business environment. A strong business ecosystem is the backbone of any successfully operating platform.

The business ecosystem is perceived as a means of offering that shapes the way participants operate and facilitates their joint development [6]. This approach is further supported by the research of Parker, Van Alstyne, and Choudary [7], who focuses on the inherent platform feature to generate new economic opportunities. Iansiti and Levien [4] state that nowadays competition among businesses is waged at ecosystem rather than company level. This factor is the driving force behind PropTech Bulgaria's unwavering strategy to build the most complex business model archetype, namely business ecosystem.

The presence of the so-called network effect is of prime importance, as it determines that the platform value accumulates with the number of active participants [8]. The network effect is also defined as economies of scope on the demand side [2], meaning 'the value of a product or service is in function to the number of people using it' [2]. Hence, it is of uttermost importance for a platform to be purpose-built in a way as to encourage participation, interaction, and collaboration in a purposeful manner, recognizing the diverse ways to achieve a network effect to directly pivot on the level of strength of separate stakeholders. The degree of strength is reckoned as to decide on the precise structure of the platform. Eisenmann et al. [3] define factors such as subsidizing the right group of participants, avoiding price wars through platform differentiation, and being particularly cautious of larger rivals coming from unfamiliar industries who may offer the platform's service as part of their larger service packages, thus becoming more attractive to the potential clients of the platform.

PropTech Bulgaria's e-platform seeproptech.eu is established to be an "interface for business ecosystem interaction," but it does not imply co-creation of products; nor does it imply the inclusion of as many stakeholders as possible, given that it provides limited and closed access. At the same time, the e-platform of PropTech Bulgaria also presupposes that a network effect, a typical characteristic of technological platforms, should be amplified. Refer to point 5 below on our synopsis how PropTech Bulgaria's platform is technologically structured to optimize and maximize the network effect.

Baldwin & Woodard [9] assert that platforms exist in a number of different settings: within an organization as product lines; among companies as multi-product systems; or as multi-sided markets

where different types of participants interact through a common infrastructure. From the standpoint of PropTech, a rapidly developing sector at the intersection of digital technologies and the construction industry, platforms should be fitted for bringing together numerous stakeholders: technology startups, construction companies, construction investors, property developers, municipal and state institutions, universities, venture capital funds and other external funding organizations, national PropTech associations, international PropTech organizations, as well as diverse clients outside the construction industry such as banks, energy distribution companies, water and sewage firms, manufacturing businesses, etc. Such a wide range of stakeholders calls for a new type of communication platform which, unlike e-commerce platforms, is not solely transaction-oriented but prioritizes information exchange, strategic collaboration, community building and engagement, as well as social bonding. Despite the growing interest in platform business models, scientific research related to closed communication platforms like the one of PropTech Bulgaria remains limited. Jacobides, Cennamo, and Gawer [10] view such a deficiency as a critical challenge throughout the process of constructing resilient digital ecosystems. A substantial gap in the scientific conceptualization of business models in such platforms is observed, exacerbated by the lack of research on determining appropriate business and software architecture and methodology; and evaluating the network effect in the absence of open access.

3. Objective, Research Approach, and Conceptual Model

The purpose of this research is to practically apply and illustrate how PropTech Bulgaria has been construing the organization's business ecosystem on the grounds of the Business Models Archetypes instrument of Cabage [11]. Showcasing the process of actual application, we exemplify how the design and implementation of a network of e-platforms is being steadily directed within PropTech Bulgaria. This real business case aims at proving the rationale behind consolidating and structuring different business models in a mutually amplifying way, thus building up to an ecosystem, the most complex business model. The approach is firmly grounded on the paradigm that a successful ecosystem cannot emerge spontaneously, unless one mandatory prerequisite has been put to effect, namely, there is a pre-established stable foundation of separate operational business models, each technologically implemented through an e-platform. At PropTech Bulgaria we observe a logical and sequential upgrade starting from basic through combined business models, thus setting the stage for building an integral ecosystem.

3.1. Conceptual Model for Ecosystem Building through a Network of E-Platforms

The conceptual instrument includes three levels of business models, graded by their degree of complexity and integration. Figure 1 exhibits the explicit application of the Business Models Archetypes of Cabage [11] to the online presence and e-platforms of PropTech Bulgaria, thus illustrating the fact that the company has systematically and steadily been enhancing the digitalization of its services/products for more than 9 years to the date of our research, in order to prepare for establishing an ecosystem. Table 1 below tracks the development of all websites and e-platforms via a time continuum, starting from 2016 till now (August 2025).

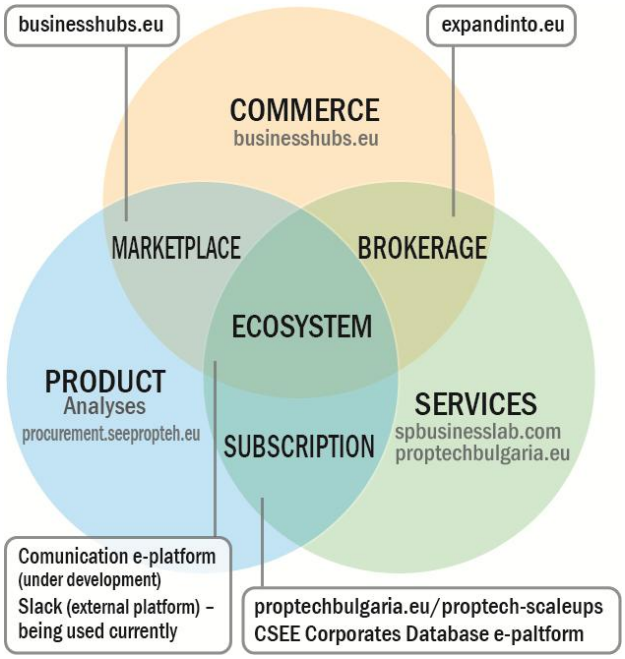


Figure 1. Conceptual Model of PropTech Bulgaria.

3.2. Basic business models of PropTech Bulgaria

Basic business models of PropTech Bulgaria and corresponding e-platforms are as follows:

- **Product Business Model:** offering international analyses and reports – procurement.seeproptech.eu (the platform is built up and is active);
- **Service Business Model**: consultancy activities, mentoring entrepreneurs and senior managers in strategic business development - spbusinesslab.com, proptechbulgaria.eu (websites are built up and are active);
- **Commerce Business Model**: online sale of products and services – businesshubs.eu (e-commerce platform is built up and is active).

3.3. Combined Business Models and Platforms

Subscription Business Model (Product + Service): access to consultations and analyses through annual membership in the association – prexcelerate.eu, proptechbulgaria.eu/ proptech-scaleups (websites are built up and are active). CSEE Corporates Database, the newest e-platform of PropTech Bulgaria, also has a Subscription business model – seecorporatedatabase.seeproptech.eu (analytical platform for corporate decision-making – built up and active).

Marketplace Business Model (Product + Commerce): offering services in a closed external network (Slack) and future public integration of Slack groups (external platform); upcoming implementation of access.seeproptech.eu (communication platform under development).

Brokerage/Intermediation Business Model (Service + Commerce): international business development services applying commission payments - expandinto.eu (website is built and active).

3.4. Ecosystem Level – Integration Platform

Communication Platform: central hub for coordination, collaboration, and resource exchange among participants from different countries – acess.seeproptech.eu (under development).

E-platform CSEE Corporates Database consists of a large database of construction investors from Central-Southeast Europe; built-up functionalities for categorization, search, and results visualization, thus digitizing the management decision-making processes required at strategizing for international expansion to other countries - seecorporatedatabase.seeproptech.eu (the analytical platform for managerial decision-making is built up and active).

This network of e-platforms is invented to not only support various business interactions but to generate a network effect which amplifies the value of both the platforms, and the entire ecosystem. The transition from single, self-sufficient solutions to an integrated environment emanates from the strategic transformation of PropTech Bulgaria's digital presence towards resilient and scalable international development.

4. Platform Architecture and Software Implementation

Building a network of e-platforms within PropTech Bulgaria requires a carefully designed architecture that ensures both the autonomy of each individual platform and the technological connectivity of all platforms within a common ecosystem. The goal is to achieve a modular, scalable, and integrable system capable of supporting various business models and dynamically adapting to the needs of participants.

4.1. Architectural Model and Integration Layers

The enhancement of websites and e-platforms of PropTech Bulgaria is grounded on a marketing approach. That is, the technological profile of a given web site or e-platform is entirely structured based on the products/services offered to the respective client categories.

Table 1 tracks down the development of all websites and e-platforms of PropTech Bulgaria along a time continuum which covers a period of over 9 calendar years, starting from 2016 till now (August 2025). It also presents the type of each website/e-platform based on Chaffey's classification [12] in addition to showcasing how PropTech Bulgaria elevates its online presence and digitalization of services.

Table 1. Development of PropTech Bulgaria's online presence and OVP (online value proposition) from flagship websites to e-business platforms.

Website Launch year	Web site Type	Web site Domain
2016 - 2018	Flagship web sites	spbusinesslab.com *(active)* impactventures.today *(deactivated)*
Upgraded to		
2019 - 2020	Brand-building web sites as well as Service-oriented web sites [12]	prexcelerate.eu *(active)* proptechbulgaria.eu *(active)*
Upgraded to		
2020	E-commerce web site, as well as communication e-platform of a social network type (social network or community building) [12]	businesshubs.eu *(active)*
2021	Seeproptech.eu as a brand-building & service-oriented web site – (1st level), but with the potential to be upgraded to e-business e-platform (2nd level). The 2nd level has already been launched in 2024 - 'Procurement' functionalities and further expanded in July 2025 – 'CSEE Corporates Database'.	seeproptech.eu *(active)* Till 07.2024 г. this web site is limited to brand-building & service-oriented one.
2022	Expandinto.eu as a brand-building и service-oriented web site [12]	expandinto.eu *(active)*

Upgraded to		
2022 - 2024	E-platform which digitizes business processes (e-business web site and platform) [12]	seeproptech.eu *(active)* Following 07.2024 г. it has become an e-business web site as 'Procurement' functionalities were launched. procurement.seeproptech.eu *(active)*
Upgraded to		
2025	2 out of 3 more modules are under development. These are additional to seeproptech.eu, thus making this e-business platform an accomplished ecosystem. The 1st module is the analytical e-platform 'CSEE Corporates Database', which was launched and has been fully operational since 15 June 2025. seecorporatedatabase.seeproptech.eu *(active)*	

Base Layer – Functional Modules. Each website (e-platform) represents an independent functional module implementing a specific business model which is depicted in Figure 1. These modules operate independently and apply different technologies as described in details in point 4.2. below.

Upper Layer – Collaboration and Communication Hub. This is access.seeproptech.eu e-platform, which will function as a central co-operation hub streamlining the international interactions within the ecosystem. The platform will integrate functions which enable closed internal communication, resource access, partner mediation and brokerage, and service marketplace. It is planned to be built on a closed SaaS infrastructure, with role-based access management and encapsulated communication channels (forum, chat, direct links to other platforms).

4.2. Technology Stack

Each e-platform uses a specific set of technologies simultaneously tailored to both its functionality, and the target client segments of PropTech Bulgaria:

- **spbusinesslab.com** – a flagship web site built with HTML, CSS, and JavaScript;
- **proptechbulgaria.eu** – a brand-building web site implemented with WordPress;
- **businesshubs.eu** – an e-commerce platform built on WooCommerce with integration to a plugin providing payment services in four currencies, including currency conversion, as well as a product catalogue plugin;
- **prexcelerate.eu** – a service-oriented web site created with WordPress, allowing for subscription form submissions;
- **expandinto.eu** – built with WordPress as a brokerage platform with registration forms;
- **procurement.seeproptech.eu** – an e-platform implemented as a tool for searching and filtering suppliers, with specific functionality for company classification and business inquiry management. The technical implementation uses HTML, CSS, JavaScript, and an external email service;
- **access.seeproptech.eu** – in the design phase; planned to be implemented using Spring Boot, Java, and Thymeleaf;
- **CSEE corporate database** – an e-platform built with Spring Boot, Java 17, and Thymeleaf for creating a dynamic web interface. The platform uses Spring Data JPA and MySQL for managing and accessing a corporate database, supporting CRUD operations, filtering, and searching database records. Input data validation is carried out using spring-boot-starter-validation, ensuring reliability and security during the process of company registration. RESTful APIs are integrated for external communication with other systems, and spring-boot-starter-security for controlling access to sensitive information. The visualization of statistical data uses the external library Chart.js. The platform is designed for administration, search, corporate and management analyses of construction investors and property developers in Central-Southeast European region.

4.3. Interaction among platforms

Communication among separate modules is defined through pre-established business processes that allow for:

- Automatic user redirection between web sites/e-platforms via URL logic and SSO;
- Transfer of requests from one platform to another (e.g., from consultation to commerce or vice versa);
- Shared visual elements in graphic, web, and motion design, as well as shared structures to ensure user consistency.

4.4. Closed communication platform implementation peculiarities

Access.seeproptech.eu e-platform, being a central component of the future ecosystem, demands for a specific approach:

- The system must maintain restricted access while simultaneously providing communication among participants from different countries;
- Building a flexible backend for managing dynamic content and services.

4.5. Contribution of the architectural model to digital ecosystems' implementation

The architectural model presented does not only provide a functional foundation for building a network of e-platforms. Furthermore, it displays a conceptual framework for fostering digital business ecosystems. Such a multi-faced contribution is manifested in a number of aspects:

- **Integration of different business models.** Each of the primary business models (product, service, commerce, subscription, marketplace, and brokerage) is accomplished through a separate e-platform - an approach which allows for flexible value management within the entire system.
- Though not being technologically connected, all platforms serve a **common strategic goal** and are construed on a unified modular fundamental, allowing for comprehensive value management within the ecosystem boundaries.
- **Step-by-step development of a network effect.** The model speeds up the gradual expansion of the ecosystem through a sequential integration of participants and platforms, thus installing conditions for organic growth and involvement of international partners. Given that PropTech Bulgaria's platform is of a multi-sided type, its structure is modular. Consequently, transactions between different parties are controlled by the platform owner in the most direct way. Although modularity increases competition intensity, this structure also simplifies transactions, reducing the time parties need to reach an agreement [13]. In short, modularity helps minimize the time otherwise needed by participants to conclude a transaction. Each module has a separate network effect, meaning that in each case, the subsidized side used to attract the paying side is different. For example, in the communication platform case, the subsidized side is the national PropTech associations, which hold the direct contact with the largest construction investors in their country, while the paying side is the PropTech companies that pay to use the associations' services to reach construction investors in the respective country. Hence, the more associations there are on the platform, the greater the number of countries that PropTech companies have access to.

Another characteristic of PropTech Bulgaria's platform that impacts the network effect growth is the extent of differentiation the platform has in terms of functionalities, incoming and input information, and information exchange [3]. Differentiation immediately changes the focus of competition among rival platforms from price-based to ecosystem scope-based, therefore, putting the stress on the ecosystem that serves as the backbone of the platform.

Formalizing organizational digital structure. The systematic approach to architecture formalizes the rationale behind business transformations in PropTech Bulgaria.

Potential for standardization. The architectural model exhibits the potential for developing a reference standard on which ground similar organizations can design their ecosystems on a modular, API-oriented, and communication-active infrastructure. As Eisenmann et al. [3] underline, the market for multi-sided platforms in certain sectors is likely to be dominated by only one or at most two platforms. This is precisely the case of international PropTech mediation (the sector which PropTech Bulgaria operates in). Accordingly, establishing an industry standard platform in this sector will determine the dynamics of competition.

5. Discussion, Contributions, and Target Users Outreach Analysis

The development of a network e-platform architecture within PropTech Bulgaria represents a strategic and methodological transition from relatively independent digital initiatives to a comprehensive ecosystem framework. Not only has this transition its origin in a technological transformation but it results from reorganizing business logic in a way to align with the principles of interconnectedness, integration, and collaboration. The real business case of PropTech Bulgaria clearly demonstrates how any change in an organization's business modeling promptly impacts its technological presence. A major contribution of the developed conceptual model is the formalization of the relationship among different business model types and the corresponding e-platforms, with a clear architectural and functional role for each. A decentralized structure is typical for the early stages of digital platform development, as noted by de Reuver, Sørensen, and Basole [14]. The current model demonstrates how their functions complement and build upon each other, thus amplifying the network effect and new value creation for priority stakeholders.

5.1. Effectiveness of the Integrated Model

The implementation of a network of e-platforms allows for:

- Addressing different user needs – from B2C services (in this case a company's clients are individual end-users) to multi-facet B2B

international business interactions (in this case a company's clients are other businesses);
- Ensuring a logical transition when upgrading between different business models – for example, from subscription to direct participation in a marketplace;
- Tracking and analyzing user behavior at various touchpoints;
- Supporting business processes with varying degrees of complexity and engagement – from purchasing a product to strategic partnerships.

In this context, the network structure creates synergy between the platforms which does not otherwise exist if they are used as stand-alone units.

5.2. Reaching Target Users

The analysis of online behavior and the dynamics of single platforms usage shows that:

- The upcoming platform access.seeproptech.eu has the potential to serve as a catalyst for the actual community building potential, taking advantage of the deep interaction among participants from different countries.
- The analysis of current development indicates that the architecture is suitable for progressive but simultaneous user engagement towards achieving a higher degree of integration and participation.
- The analysis of current information consumption proves that the modular structure significantly increases the speed of interaction among different participants on PropTech Bulgaria's platform, thus reducing the time needed for transactions between them or between them and PropTech Bulgaria to be completed.
- The analysis of current communication on the platform also proves that the subsidized parties in each module are correctly specified, as the unfolding of a network effect is being observed in each module.

6. Conclusion and Further Research

This study presented a conceptual model for building a business ecosystem through design and implementation of a network of e-platforms that integrate software services within PropTech Bulgaria. The approach is based on an in-depth analysis of the organization's business logic, systematic mapping of its platforms, and constructing a software architecture that allows both autonomy and connectivity between individual modules. The core of the established ecosystem comprises a hierarchy of business models, categorized by their degree of integration – from basic forms such as "Product," "Service," and "Commerce" to more complex hybrid models like "Subscription," "Marketplace," and "Brokerage." A specific e-platform is the technological embodiment of each of these business models, and supports the respective business model to play its role in achieving the strategic goal – building a resilient, international PropTech community. The network architecture is designed to provide for expansion and adaptation, based on modularity, API connectivity, and flexible communication logic. The upcoming implementation of the closed communication platform access.seeproptech.eu will introduce the last missing functionality – centralized coordination and community building. Although the system has not yet reached its full capacity, the analysis shows a high degree of readiness and steady deployment of a network effect. The conceptual model under consideration demonstrates not only how platforms can be projected and connected, but furthermore, it reveals how business models can be managed through a software architecture oriented towards value, synergy, and co-development.

As a next step in the research process, it would be very useful to apply the model of Rong et al. [15], which would further visually illustrate and explain how and why the web platforms built by PropTech Bulgaria change over time; how such a change (first dimension) translates into a change in customer segments and the company's value proposition (second dimension) given the change in the life cycle the company is in (third dimension).

Based on the current study, several important directions for development can be outlined:

- Technical completion and launch of access.seeproptech.eu with role-based access management, communication modules, and related services;
- Expansion of interfaces between platforms through SSO and unified UX elements;
- International expansion of the partner base by introducing new participants;
- Formalization of the model into an open standard that serves as a reference framework for building e-business ecosystems in similar contexts.

References

[1] Zhao, Y., von Delft, S., Morgan-Thomas, A., & Buck, T. (2020). The evolution of platform business models: Exploring competitive battles in the world of platforms. *Long Range Planning*, *53*(4), 101892. https://doi.org/10.1016/j.lrp.2019.101892

[2] Kim, J. (2014). Platform business and network strategy. *STI Policy Review*, *5*(1), 57–74.

[3] Eisenmann, T., Parker, G., & Van Alstyne, M. (2006, October). Strategies for two-sided markets. *Harvard Business Review*.

[4] Iansiti, M., & Levien, R. (2004). *The keystone advantage: What the new dynamics of business ecosystems mean for strategy, innovation, and sustainability*. Harvard Business School Press.

[5] Cusumano, M. (2010). Technology strategy and management: The evolution of platform thinking. *Communications of the ACM, 53*(1), 32–34. https://doi.org/10.1145/1629175.1629189

[6] Moore, J. (1993). Predators and prey: A new ecology of competition. *Harvard Business Review, 71*(3), 75–86.

[7] Parker, G., Van Alstyne, M., & Choudary, S. P. (2016). *Platform revolution: How networked markets are transforming the economy and how to make them work for you*. W. W. Norton & Company.

[8] Evans, D., & Schmalensee, R. (2016). *Matchmakers: The new economics of multisided platforms*. Harvard Business Review Press.

[9] Baldwin, C. Y., & Woodard, C. J. (2009). The architecture of platforms: A unified view. In A. Gawer (Ed.), *Platforms, markets and innovation* (pp. 19–44). Edward Elgar Publishing. https://doi.org/10.4337/9781849803311.00008

[10] Jacobides, M. G., Cennamo, C., & Gawer, A. (2018). Towards a theory of ecosystems. *Strategic Management Journal, 39*(8), 2255–2276. https://doi.org/10.1002/smj.2904

[11] Cabage, N., & Zhang, S. (2013). *The smarter startup*. Retrieved from https://nealcabage.com/framework/business-model-archetypes/

[12] Chaffey, D., & Smith, P. R. (2008). *eMarketing eXcellence: Planning and optimizing your digital marketing* (3rd ed.). Butterworth-Heinemann/Elsevier. https://doi.org/10.4324/9780203082812

[13] Baldwin, C. Y. (2007). Where do transactions come from? Modularity, transactions, and the boundaries of firms. *Industrial and Corporate Change, 17*(1), 155–195. https://doi.org/10.1093/icc/dtm036

[14] De Reuver, M., Sørensen, C., & Basole, R. C. (2018). The digital platform: A research agenda. *Journal of Information Technology, 33*(2), 124–135. https://doi.org/10.1057/s41265-016-0033-3

[15] Rong, K., Lin, Y., Shi, Y., & Yu, J. (2013). Linking business ecosystem lifecycle with platform strategy: A triple view of technology, application and organization. *International Journal of Technology Management, 62*(1), 75–94. https://doi.org/10.1504/ijtm.2013.053042

Computer Science and Interdisciplinary Research Journal
Volume 2, Issue 2, 2025. ISSN: 3033-1218, DOI: 10.70862/CSIR.2025.0202-15

Cloud-Based Educational Transformation: The Municipal Policy Model of Plovdiv, Bulgaria

Stefan Stoyanov

University of Plovdiv Paisii Hilendarski, 236, Bulgaria Blvd, 4027 Plovdiv, Bulgaria

stefan.stoyanov@uni-plovdiv.bg

Abstract – The article presents the process of digitalization of education in the Municipality of Plovdiv, Bulgaria, as a good practice for educational transformation. Based on a comparative analysis of international practices and the experience of the Municipality of Plovdiv, a model for the digital transformation of the school environment has been developed. The paper examines the main stages of the implementation of the "School in the Cloud" project, which includes the activation of cloud platforms, the training and certification of teachers, as well as the application of the 1:1 model. The achieved results are presented, highlighting the region as one of the leaders in Eastern Europe in implementing cloud technologies in education. The study emphasizes the importance of local policies, teacher training and professional development, as well as the need for active community engagement. The establishment of effective cooperation between local authorities, schools, and technology partners stands out as a prerequisite for the sustainable development of the educational environment in the digital era. The described strategies can serve as a model for the modernization and digitalization of municipal-level educational systems.

Keywords – cloud technologies, digital transformation, educational policy, municipal education, Google Reference School, teacher training.

DOI: https://doi.org/10.70862/CSIR.2025.0202-15

Corresponding author: Stefan Stoyanov, University of Plovdiv Paisii Hilendarski
Email: stefan.stoyanov@uni-plovdiv.bg

Received: 24 October 2025.
Revised: 6 November 2025.
Accepted: 8 November 2025.
Published: 8 November 2025.

© 2025 Stefan Stoyanov. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

The article is published with Open Access at http://csirjournal.com/

1. Introduction

New technologies have the potential to redefine and transform educational systems, as well as the methods of teaching, learning, and assessment. Cloud computing provides a wide range of tools that enable innovation in the organization of the learning process and in the communication between participants in the educational system. Despite the significant potential of cloud technologies, their implementation at the level of individual educational institutions is associated with a number of challenges. Among these are the need for qualified teaching staff, the adaptation of curricula, the provision of appropriate resources, and the establishment of reliable infrastructure. In this context, the role of municipal policies that govern local educational systems becomes particularly important. The good practice of the Municipality of Plovdiv in developing and implementing a strategy for the adoption of cloud solutions serves as an example worthy of attention both nationally and internationally.

The study presented in this article focuses on the analysis of municipal policies in Bulgaria concerning school education and their role in promoting the implementation of innovations and technologies in teaching and learning. It examines the process through which municipalities formulate their educational strategies and the extent to which these align with national priorities for digitalization and for improving students' preparedness for the labor market. The challenges faced by schools in implementing cloud-based solutions are analyzed, along with the corresponding measures taken to overcome them. Concrete examples from the practice of the Municipality of Plovdiv illustrate the effects of these policies and their applicability at the local level. The data used for the local educational ecosystem was collected from official reports of the Plovdiv Municipality and decisions of the local parliament.

2. Modern Educational Policies

Modern educational policy can be defined as a set of principles, guidelines, and regulations that determine how education functions within a given country or region. It covers aspects such as institutional management, curriculum design, teacher qualifications, financing, and student assessment. Its primary goal is to ensure access to quality education for all students, supporting the development of their potential and their active participation in society.

The formation of educational policy depends on multiple factors, including public attitudes, economic conditions, political decisions, and global educational trends. Policy frameworks are often based on scientific and ethical principles, encompassing the right to education, equality of opportunity, and the need to develop competencies that meet the requirements of the modern labor market. In this context, municipal educational policy represents a localized application of national policy. It is developed and implemented by local authorities in accordance with the educational needs of the specific community. Municipalities manage funding, the number and type of schools, the provision of qualified teachers, and infrastructure renewal. Differences in access to resources and in the challenges faced determine the specific nature of local educational strategies.

In the 21st century, education systems worldwide face global challenges – foremost among them the rapid technological progress transforming how knowledge is delivered and acquired. Institutions must adapt teaching and assessment methods while emphasizing digital literacy [1–4]. This trend became particularly evident during the pandemic, which led to the widespread adoption of distance learning [5,6].

On the other hand, demographic changes and migration flows have a direct impact on the composition of student populations, creating challenges for the integration of children from diverse cultural and linguistic backgrounds. This necessitates innovation in pedagogical approaches and the application of personalized learning in order to address the growing diversity of educational needs. In this process, digital technologies play a key role. The integration of artificial intelligence, virtual reality, and augmented reality offers new opportunities to enrich the learning process, making it more engaging and adaptable to the individual needs of students.

Despite these promising prospects, funding remains one of the most serious challenges facing educational systems. Insufficient budgeting in most municipalities and the need for larger national funds limit the opportunities for modernizing infrastructure and for providing the necessary qualification of teaching staff capable of implementing innovative practices.

Another critical aspect is the deficit in systems for assessing and monitoring educational outcomes. Existing standards often fail to adequately reflect students' actual skills and knowledge, as well as the achievements of teachers and schools when considered within the context of their specific conditions and available resources.

The sustainability of educational systems largely depends on the professional preparation and development of teachers. In an era when continuous learning is essential, it is important that teachers receive adequate training and opportunities for professional growth, enabling them to respond effectively to rapidly changing working conditions [7].

Modern global competition imposes even higher demands on education. It must prepare not only competent professionals but also innovators capable of critical thinking and creativity. This determines the need for curricular reforms that emphasize digitalization and the development of "soft" skills such as problem-solving, critical analysis, and creative thinking. The application of educational technologies aligned with the realities of the digital world is one of the key approaches to successfully addressing the challenges of the 21st century [8].

A comparative analysis of educational policies in Bulgaria – and particularly in Plovdiv – with practices from other educational systems shows that there are numerous models that can be adapted to the national and local context. As part of the global educational environment, Bulgarian schools need to align their strategies to meet contemporary challenges and take advantage of the opportunities offered by new technologies and teaching methods.

Some key findings from the conducted analysis are as follows:

- Bulgarian schools can adopt successful practices from the United States and the Scandinavian countries, where emphasis is placed on the individualization and diversity of curricula. This can be achieved through the integration of modern technological solutions, including cloud technologies [9-11].
- Singapore and South Korea provide examples of systems with high academic standards and excellent performance on international educational assessments. Despite their traditionally rigorous approach, in recent years they have also focused on pedagogical innovation and technological integration, which foster students' creativity and critical thinking alongside knowledge retention [12,13].

- Many countries are developing programs for inclusive education and the integration of students from vulnerable social groups. Successful practices in the United Kingdom and Germany can be used to enhance Bulgarian policies by building stronger support systems, often based on digital solutions [14,15].
- The experience of the United Kingdom and Finland demonstrates that the successful implementation of educational policies requires the active involvement of local authorities, parents, businesses, and non-governmental organizations. Such collaboration improves the quality of education by providing additional resources and expert support [16,17].
- Teacher support and professional training are key factors for successful digital transformation. In this context, the example of Finland, where the emphasis is placed on systematic teacher training and qualification, is particularly illustrative and can serve as a model for Bulgarian municipalities [18].

The findings are summarized in Table 1.

Table 1. International Good Practices and Their Applicability to Bulgaria and the Municipality of Plovdiv.

Country / Region	**Key Practices**	**Potential Applicability for Bulgaria / Plovdiv**
United States, Scandinavian countries	Individualization and diversity in curricula.	Introduction of personalized learning through cloud technologies and digital resources.
Singapore, South Korea	High academic standards combined with technological innovation.	Development of creativity and critical thinking through the integration of traditional knowledge and new technologies.
United Kingdom, Germany	Inclusive education and integration of vulnerable groups.	Building a stronger support system through digital tools and cloud-based solutions.
United Kingdom, Finland	Active involvement of the community (local authorities, parents, business, NGOs).	Improving quality through partnerships and the mobilization of local resources.
Finland	Support and systematic teacher training.	Qualification for the use of new technologies and sustainable implementation of digitalization.

The educational policy of the Municipality of Plovdiv is undergoing a stage of renewal in which the implementation of innovations and technological solutions has become a necessity. With the adoption of the new educational strategy, the program for the digitalization of the learning process was launched in 2015. A comparison with global models indicates that digital transformation serves as a key mechanism for modernization and for improving the quality of education at both the local and national levels.

3. Policy of the Municipality of Plovdiv for the Implementation of Cloud Technologies in School Organizations

The Municipality of Plovdiv is the first institution in Bulgaria to develop and implement an independent digitalization policy that provides all municipal schools and kindergartens with an activated cloud platform accessible through their own domain. This achievement is the result of targeted efforts and a consistent strategy implemented over several years.

The process of digital transformation in education began in 2015 as a policy initiative of the Municipality of Plovdiv, in partnership with the city's schools and technological partners. The process can be divided into three main stages:

1. Providing access to cloud-based learning platforms;
2. Training of municipal experts and school teams in working with cloud platforms;
3. Implementing cloud technologies as a means of transforming the learning process.

The first stage (Figure 1) began in 2016, when the team of the Department of Education at the Municipality of Plovdiv identified an opportunity for reform aimed at the digitalization of school education. The initiative was motivated by the fact that the cloud services provided by technology companies Google and Microsoft had already been fully translated into Bulgarian, made publicly accessible, and supported a wide range of educational applications. These tools enabled both the digitalization of administrative processes in schools and their integration into the learning process, with the aim of enhancing digital literacy and modernizing teaching methods.

The main advantages that cloud technologies offer in education include:

- facilitated access to diverse educational resources and tools, regardless of the location of students and teachers;
- improved collaboration and real-time communication;
- optimization of time and organization of the learning process;
- reduction of administrative and financial costs;

- opportunity for storing, sharing, and analyzing large volumes of data [19-22].

In collaboration with the official partners of both companies, the Municipality of Plovdiv provided schools with the opportunity to choose between the two platforms. As a result, 50 schools implemented Google Workspace for Education (then known as Google G Suite), while two schools opted for Microsoft 365 Education.

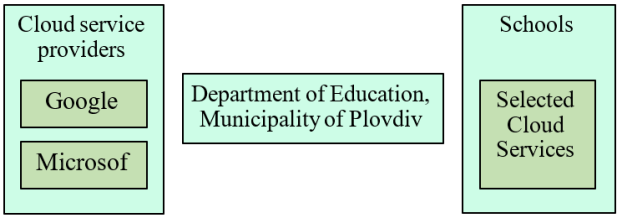


Figure 1. Access to Cloud-Based Learning Platforms.

The positive results achieved during the first stage gave both the Municipality of Plovdiv and the school teams confidence in the significance of the undertaken initiative. This led to the decision to expand the project and implement the following two stages.

The second stage (Figure 2), covering 2017, focused on establishing a sustainable mechanism to support schools in using the selected cloud platforms. A practice was introduced whereby all administrative communications between the municipality and educational institutions were conducted through official cloud-based email accounts. This measure ensured data security, centralized information management, and compliance with the General Data Protection Regulation (GDPR) requirements.

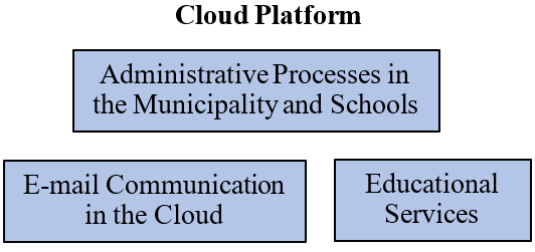


Figure 2. Introduction and Training for Working with Cloud Platforms.

To ensure the high quality of implementation of the planned policies, the team from the Department of Education was the first to undergo specialized training for working with cloud systems. During this stage, the "School in the Cloud" project [23] was launched, including the development of a strategic framework and educational programs aimed at financing and supporting schools on their path toward digital transformation.

As a result, numerous training sessions were conducted for school teams, enabling them to acquire practical skills for working with cloud platforms. This facilitated the digitalization of various administrative processes and significantly improved internal school communication.

The third stage (Figure 3) began once schools had the necessary infrastructure and trained staff, marking the start of the actual integration of cloud platforms into educational practice.

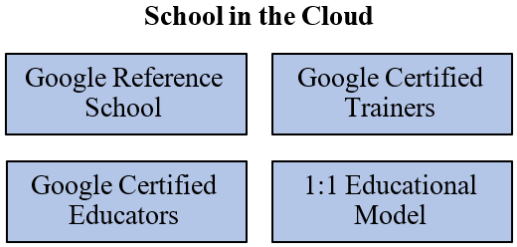


Figure 3. Elements of the Integrated "School in the Cloud" Ecosystem of the Municipality of Plovdiv.

A significant number of teachers completed certification programs under Google for Education, gaining internationally recognized competencies for working in a digital learning environment. These efforts created the foundation for introducing the innovative 1:1 educational model, which provides personal devices for all students and teachers, to be used both in the classroom and during distance learning [24,25].

The high level of engagement and the sustainability of the achieved results led to international recognition for a number of Plovdiv schools. Among them are Ivan Vazov Language High School, Chernorizets Hrabar Secondary School, Saints Cyril and Methodius Secondary School, Hristo G. Danov Secondary School, Vocational School of Food Technology and Engineering, Yane Sandanski Primary School, and Exarch Antim I Primary School, all of which received the status of Google Reference School.

In a parallel project implemented with the support of the Municipality of Plovdiv, teachers were certified under the international Google for Education programs (Google Certified Educator – Level 1 and Level 2) [26, 27].

To support this process, the Municipality of Plovdiv established two new programs – "Digitalization of the Learning Process" and "Modernization of the Learning Process" – through which schools were assisted in implementing cloud platforms and upgrading hardware during the period 2017–2019. These programs provided targeted funding for:

- initial activation and configuration of the cloud platform within the school's domain;
- training of teachers, principals, and administrators;
- purchase of devices and additional equipment, co-financed at 50–70% of the total cost.

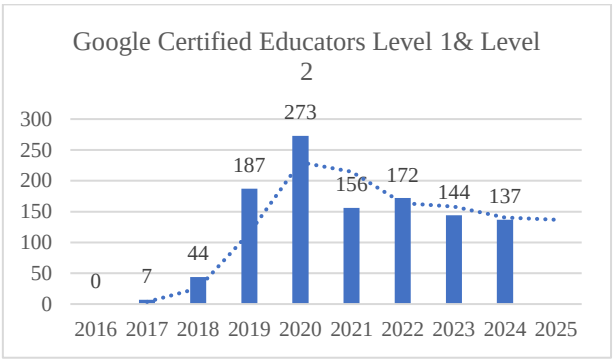
4. Results

The key results from the implementation of the education digitalization projects in the Municipality of Plovdiv are summarized in Table 2:

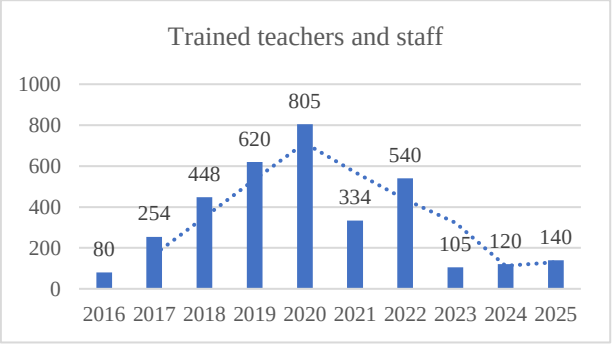
- By 2019, three years after the launch of the "School in the Cloud" project, 52 municipal and 4 state and private schools, as well as the Department of Education of the Municipality of Plovdiv, had fully functioning cloud platforms [28].
- Each year, numerous teachers and administrators participate in training programs.
- The number of Google Certified Trainers increased from 2 in 2019 to 7 in 2020.

Table 2. Dynamics of Cloud Technology Implementation and the Related Educational Transformation in the Municipality of Plovdiv (2016–2025).

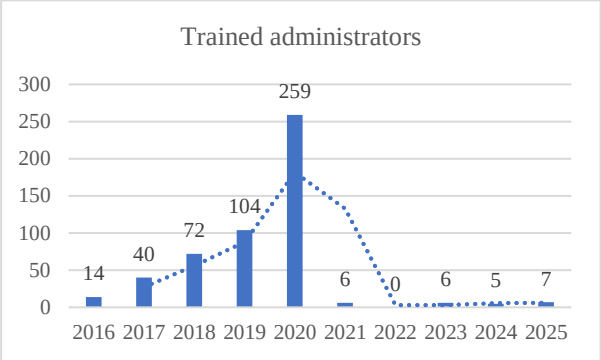
Indicator \ Year	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
Schools with active cloud services	7	20	52	52	52	52	52	52	52	52
Trained teachers and staff	80	254	448	620	805	334	540	105	120	140
Trained administrators	14	40	72	104	259	6	10	6	5	7
Google Certified Educators (Level 1 & Level 2)		7	44	187	273	156	172	144	137	
Google Certified Trainers				2	7	1	2			
Classes under the 1:1 model			1	6	21	22	38	94	95	69
Google Reference Schools				1		5	1			



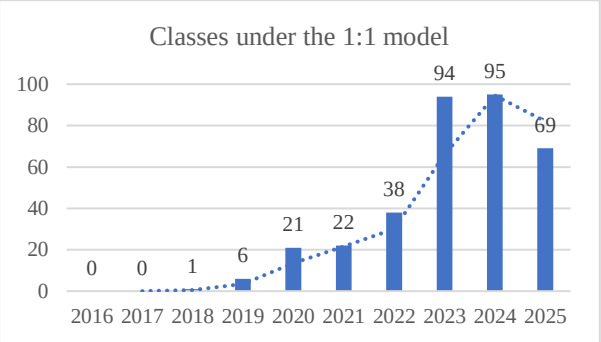
(a)



(b)



(c)



(d)

Figure 4. Trends for: **(a)** Google Certified Educators Level 1& Level 2; **(b)** Trained teachers and staff; **(c)** Trained administrators; **(d)** Classes under the 1:1 model.

Trends for certified educators, administrators, trained teachers and staff, as well as for classes under the 1:1 model are presented in Figure 4.

Particular attention should be paid to several key achievements related to the implementation of educational digital transformation policies in the Municipality of Plovdiv:

- The adoption of the 1:1 educational model in a significant number of schools, ensuring equal access to digital devices and creating conditions for personalized learning;
- The certification of seven schools as Google Reference Schools, a recognition that confirms the high level of integration of cloud technologies and pedagogical innovations;
- The establishment of the first officially recognized independent community of teachers in Bulgaria – GEG Plovdiv (Google Educators' Group), which functions as a platform for mutual assistance, technological collaboration, and methodological support among educators.

4.1. The 1:1 Educational Model

Plovdiv has established itself as a national leader in Bulgaria, with the highest number of teachers certified under the international Google for Education programs – a total of 273 educators by 2020. This achievement represents a natural foundation for the next phase of digital transformation – the preparation

of Google Certified Trainers and the implementation of the 1:1 educational model.

Under this model, students and teachers:

- have personal laptops and constant internet access;
- learn, collaborate, and communicate in teams both inside and outside the classroom;
- create their own digital products, applying knowledge and skills from various subjects;
- use screen time meaningfully and productively, mastering modern digital tools;
- work in a safe physical and online environment;
- plan learning activities collaboratively [25,29].

The 1:1 educational model is an established approach in numerous innovative schools worldwide – in the Netherlands, Finland, Sweden, and other countries. In these nations, hundreds of municipalities already apply it as a standard component of their educational systems. Such systems traditionally occupy leading positions in rankings of digital skills, public-sector innovation, and technological business development, while their students demonstrate high levels of functional literacy [30-32].

According to data from the Department of Education of the Municipality of Plovdiv, during the 2022/2023 school year, education under the 1:1 model was implemented in 17 schools, encompassing a total of 79 classes, once again placing Plovdiv in a leading position nationwide in the application of this approach.

4.2. Google Reference Schools

By mid-2022, there were 16 Google Reference Schools operating in Bulgaria – the only ones of this type in Central and Eastern Europe. Of these, seven are located within the city of Plovdiv, which represents a remarkable achievement and a direct result of the municipality's consistent and strategic local policy for the digitalization of education.

Google Reference Schools are institutions recognized by Google for their outstanding use of technology to promote positive learning outcomes. Schools holding this status commit to sharing their experiences and best practices, welcoming visitors, and hosting training sessions and educational events [33].

4.3. Google Educators' Group Plovdiv

In Plovdiv, the first independent teacher community in Bulgaria officially recognized by Google was established, aimed at mutual support and both technological and methodological assistance. The Google Educators' Group (GEG Plovdiv) was created by teachers, for teachers and operates on the principles of voluntarism and open access. Membership is free and voluntary, with the primary goal of collaboratively exploring new approaches to organizing the learning process through digital technologies, as well as actively sharing accumulated experience.

The community's activities include monthly methodological meetings, training events, and seminars dedicated to pedagogical innovation and the practical application of cloud technologies in education.

5. Conclusion

The use of modern technologies has become an integral part of improving the quality of education. The integration of cloud technologies into the school system of the Municipality of Plovdiv has provided new opportunities for learning, collaboration, and the development of digital competences among students and teachers. Within a few years, a comprehensive model was built, combining technological infrastructure, qualified human resources, and strategic management. The process of implementing cloud technologies followed clearly defined stages – ensuring access to platforms, training pedagogical and administrative staff, and transforming the learning process through integrated digital tools.

The achieved results – certified teachers, the implementation of the 1:1 educational model, the establishment of Google Reference Schools, and the active professional community GEG Plovdiv – demonstrate the high degree of sustainability of the model. It is important to emphasize that, thanks to the digitalization policies and measures undertaken four years before the COVID-19 pandemic, the schools in Plovdiv were already prepared for the transition to distance learning and succeeded in ensuring the continuity of the educational process.

Today, education faces new horizons and challenges associated with the emergence of GenAI tools – a natural continuation of the evolution of artificial intelligence and cloud technologies. The easy access to these tools carries both enormous potential and certain risks: they can enhance creativity and analytical thinking but may also reduce independent reasoning and depth of understanding if not used consciously and ethically. Enhancing knowledge and skills for working with GenAI tools is the next strategic step that the educational community must undertake. It is necessary to develop timely policies and training programs focused on the qualitative and ethical use of artificial intelligence. The open but probabilistic nature of AI-generated knowledge requires learners to demonstrate higher levels of critical thinking, analytical ability, and self-directed learning and creativity, in collaboration with intelligent technologies [34-38].

Acknowledgements

The author expresses sincere gratitude to the project "Digital Sustainable Ecosystems – Technological Solutions and Social Models for Ecosystem Sustainability (DUECOS), BG-RRP-2.004-0001-C01", implemented within the framework of the Research Center for Innovation in Education and SMART Technologies in Learning, funded by the European Union – NextGenerationEU through the National Recovery and Resilience Plan of the Republic of Bulgaria.

References

[1] OECD. (2023). *Future of Education and Skills 2030/2040*. Organisation for Economic Co-operation and Development. https://www.oecd.org/en/about/projects/future-of-education-and-skills-2030.html

[2] van der Vlies, R. (2020). *Digital strategies in education across OECD countries: Exploring education policies on digital technologies* (OECD Education Working Papers No. 226). OECD Publishing. https://doi.org/10.1787/33dd4c26-en

[3] Timotheou, S., Miliou, O., Dimitriadis, Y., Sobrino, S. V., Giannoutsou, N., Cachia, R., ... & Ioannou, A. (2023). Impacts of digital technologies on education and factors influencing schools' digital capacity and transformation: A literature review. *Education and Information Technologies*, *28*(6), 6695–6726. https://doi.org/10.1007/s10639-022-11431-8

[4] Qayyum, A. (2022). Assessing the digital transformation of education systems: An international comparison. In *Handbook of open, distance and digital education* (pp. 1–18). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-2080-6_14

[5] Bozkurt, A., Jung, I., Xiao, J., Vladimirschi, V., Schuwer, R., Egorov, G., ... & Paskevicius, M. (2020). A global outlook to the interruption of education due to COVID-19 pandemic: Navigating in a time of uncertainty and crisis. *Asian Journal of Distance Education*, *15*(1), 1–126. https://www.asianjde.com/ojs/index.php/AsianJDE/article/view/462

[6] Antoninis, M., Alcott, B., Al Hadheri, S., April, D., Fouad Barakat, B., Barrios Rivera, M., ... & Weill, E. (2023). *Global education monitoring report 2023: Technology in education: A tool on whose terms?* UNESCO. https://doi.org/10.54676/UZQV8501

[7] Vitanova, N. (2019). Education of the future. *Pedagogy*, *91*(8), 1057–1090.

[8] Fullan, M. (2015). A New Culture of Learning: Cultivating the Imagination for a World of Constant Change. *Studies in Art Education*, *56*(3), 281–283. https://doi.org/10.1080/00393541.2015.11518969

[9] Pane, J. F., Steiner, E. D., Baird, M. D., Hamilton, L. S., & Pane, J. D. (2017). *Informing progress: Insights on personalized learning implementation and effects.* RAND Corporation. https://www.rand.org/pubs/research_reports/RR2042.html

[10] Lee, D., Huh, Y., Lin, C. Y., & Reigeluth, C. M. (2022). Personalized learning practice in US learner-centered schools. *Contemporary Educational Technology*, *14*(4), Article ep385. https://doi.org/10.30935/cedtech/12330

[11] Håkansson Lindqvist, M. (2021). The use of digital technologies in a 1:1 laptop initiative: The parent perspective. *Seminar.net: Media, Technology and Lifelong Learning*, *17*(1), 1–28. https://doi.org/10.7577/seminar.3582

[12] Tan, C., & Ng, C. S. (2021). Cultivating creativity in a high-performing education system: The example of Singapore. *Journal of Curriculum and Pedagogy*, *18*(3), 253–272. https://doi.org/10.1080/15505170.2020.1808126

[13] Chae, Y. (2021). The creativity and character in Korea teacher education. *Contemporary Educational Researches Journal*, *11*(1), 1–8. https://doi.org/10.18844/cerj.v11i1.4728

[14] Goodley, D., Cameron, D., Liddiard, K., Parry, B., Runswick-Cole, K., Whitburn, B., & Wong, M. E. (2020). Rebooting inclusive education? New technologies and disabled people. *Canadian Journal of Disability Studies*, *9*(5), 515–549. https://doi.org/10.15353/cjds.v9i5.707

[15] Knigge, M., & Kollosche, D. (2019). Inclusive education in German schools. In D. Kollosche, R. Marcone, M. Knigge, M. G. Penteado, & O. Skovsmose (Eds.), *Inclusive mathematics education* (pp. 35–47). Springer. https://doi.org/10.1007/978-3-030-11518-0_3

[16] Kiesi, I. (2023). Co-operation of edu-business and public schooling: Is the governance of education in Finland shifting from the public sector to networks? In M. Thrupp, P. Seppänen, J. Kauko, & S. Kosunen (Eds.), *Finland's famous education system: Unvarnished insights into Finnish schooling* (pp. 155–172). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-8241-5_10

[17] Greany, T. (2022). Place-based governance and leadership in decentralised school systems: Evidence from England. *Journal of Education Policy*, *37*(2), 247–268. https://doi.org/10.1080/02680939.2020.1792554

[18] Byman, R., Jyrhämä, R., Stenberg, K., Maaranen, K., Sintonen, S., & Kynäslahti, H. (2021). Finnish teacher educators' preferences for their professional development: Quantitative exploration. *European Journal of Teacher Education*, *44*(4), 432–451. https://doi.org/10.1080/02619768.2020.1793952

[19] Wang, L. Y. K. (2024). Evaluating the impact of cloud e-learning in higher education: An empirical investigation. *Knowledge Management & E-Learning*, *16*(2), 286–308. https://doi.org/10.34105/j.kmel.2024.16.014

[20] Sharma, M., Vadalkar, S., Singh, A., Chavan, G., & Tsagarakis, K. P. (2025). How Does Cloud Computing Adoption Accelerate Education? – Enhancing Academic Results by Integrating Knowledge

Management Pre and Post Pandemic. *International Journal of Human–Computer Interaction, 41*(17), 11095–11115. https://doi.org/10.1080/10447318.2024.2440673

[21] Yurchenko, A., Rozumenko, A., Rozumenko, A., Momot, R., & Semenikhina, O. (2023). Cloud technologies in education: The bibliographic review. *Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska, 13*(4), 79–84. https://doi.org/10.35784/iapgos.4421

[22] Thavi, R., Jhaveri, R., Narwane, V., Gardas, B., & Jafari Navimipour, N. (2024). Role of cloud computing technology in the education sector. *Journal of Engineering, Design and Technology, 22*(1), 182–213. https://doi.org/10.1108/JEDT-08-2021-0417

[23] Municipality of Plovdiv. (n.d.). *Specialized website "Schools in the Cloud".* https://cloud.plovdiv.bg/

[24] Learning 1:1 Bulgaria. (n.d.). *Project portal "1:1" in Bulgaria.* https://learning1to1.bg/

[25] Hall, C., & Lundin, M. (2024). Technology in the classroom: Personal computers and learning outcomes in primary school. *Economics of Education Review, 100,* Article 102536. https://doi.org/10.1016/j.econedurev.2024.102536

[26] Google for Education. (n.d.). *Fundamentals training.* https://edu.exceedlms.com/student/path/1716233

[27] Anis, M. (2024). Teacher professional development in the digital age: Addressing the evolving needs post-COVID. *International Journal for Multidisciplinary Research, 6*(1), 1–14. https://doi.org/10.4018/979-8-3373-1692-5.ch005

[28] Angelov, A. (2020). Overview of activators and activities related to integration of Google G Suite for Education in Plovdiv Municipality schools. *Education and Development, 5,* 1–12. ISSN 2603-3577.

[29] Tsokov, G., Angelov, A., Georgieva, M., Hristov, S., & Dolapchieva, S. (2020). *Digital leap in Bulgarian schools – The "One-to-One (1:1)" model.* eISBN 978-619-91278-8-9.

[30] Zheng, B., Warschauer, M., Lin, C., & Chang, C. (2016). Learning in One-to-One Laptop Environments: A Meta-Analysis and Research Synthesis. *Review of Educational Research, 86*(4), 1052–1084. https://doi.org/10.3102/0034654316628645

[31] Bergström, P., & Wiklund-Engblom, A. (2022). Who's got the power? Unpacking three typologies of teacher practice in one-to-one computing classrooms in Finland. *Computers & Education, 178,* Article 104396. https://doi.org/10.1016/j.compedu.2021.104396

[32] Cameron, D. L., Herlofsen, C., Jensen, C. R., Lindqvist, G., Molbæk, M., Ström, K., & Sundqvist, C. (2024). The special education profession in four Nordic countries in light of evolving policies on inclusion and special needs education. *European Journal of Special Needs Education, 39*(6), 849–865. https://doi.org/10.1080/08856257.2024.2425500

[33] Google for Education. (n.d.). *Official education directory.* https://edudirectory.withgoogle.com/

[34] Taguma, M., & Frid, A. (2024). Curriculum frameworks and visualizations beyond national frameworks: Alignment with the OECD learning compass 2030 (OECD Education Working Papers No. 314). OECD Publishing. https://doi.org/10.1787/2a4bdce6-en

[35] OECD. (2025). *OECD teaching compass: Reimagining teachers as agents of curriculum changes* (OECD Education Policy Perspectives No. 123). OECD Publishing. https://doi.org/10.1787/8297a24a-en

[36] Boeskens, L., & Meyer, K. (2025). Policies for the digital transformation of school education: Evidence from the Policy Survey on School Education in the Digital Age (OECD Education Working Papers No. 328). OECD Publishing. https://doi.org/10.1787/464dab4d-en

[37] Nemorin, S., Vlachidis, A., Ayerakwa, H. M., & Andriotis, P. (2023). AI hyped? A horizon scan of discourse on artificial intelligence in education (AIED) and development. *Learning, Media and Technology, 48*(1), 38–51. https://doi.org/10.1080/17439884.2022.2095568

[38] Liu, R., Zenke, C., Liu, C., Holmes, A., Thornton, P., & Malan, D. J. (2024, March). Teaching CS50 with AI: Leveraging generative artificial intelligence in computer science education. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education* (Vol. 1, pp. 750–756). ACM. https://doi.org/10.1145/3626252.3630938

Computer Science and Interdisciplinary Research Journal
Volume 2, Issue 2, 2025. ISSN: 3033-1218, DOI: 10.70862/CSIR.2025.0202-16

Integrating Economic Indicators and Machine Learning for Exchange Rate Forecasting

Ivaylo Boyoukliev [1], Vanya Ivanova [2], Hristina Kulina [3], Boyan Zlatanov [4]

University of Plovdiv Paisii Hilendarski, 236, Bulgaria Blvd, 4027 Plovdiv, Bulgaria

[1] boyouklievi@uni-plovdiv.bg, [2] vantod@uni-plovdiv.bg, [3] kulina@uni-plovdiv.bg, [4] bobbyz@uni-plovdiv.bg

Abstract – This study presents an integrated framework for forecasting the USD/JPY exchange rate using ensemble machine learning techniques – CART-Ensembles and Arcing. The analysis employs monthly macroeconomic data from October 1987 to April 2025. Key predictive variables include differentials in interest rates, inflation, unemployment, and government bond yields between the United States and Japan. After a preliminary time-series analysis and feature selection, ensemble models were developed for four distinct time periods. Each model performs step-by-step forecasts up to four months ahead (January–April 2025), achieving exceptionally high predictive accuracy, with coefficients of determination (R^2) above 0.98 and mean absolute percentage errors (MAPE) below 1%. The results demonstrate that combining macroeconomic fundamentals with advanced ensemble learning methods significantly enhances the accuracy, robustness, and reliability of exchange rate forecasting. These findings provide policymakers and investors with a robust analytical tool for anticipating short-term currency fluctuations under varying macroeconomic conditions.

Keywords – Ensemble Machine Learning Methods, Macroeconomic Indicators, Exchange Rate Forecasting

DOI: https://doi.org/10.70862/CSIR.2025.0202-16

Corresponding author: Hristina Kulina,
University of Plovdiv Paisii Hilendarski
Email: kulina@uni-plovdiv.bg

Received: 13 October 2025.
Revised: 8 November 2025.
Accepted: 8 November 2025.
Published: 9 November 2025.

© 2025 Ivaylo Boyoukliev, Vanya Ivanova, Hristina Kulina, Boyan Zlatanov. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

The article is published with Open Access at http://csirjournal.com/

1. Introduction

Foreign exchange markets play a crucial role in the functioning of the global economy, as they influence international trade, capital flows, labour markets, purchasing power, and the overall national economic policy. Exchange rate dynamics are primarily driven by fundamental macroeconomic indicators, such as inflation, unemployment, GDP, and official interest rates, as well as the yields on local and international debt instruments. These interactions produce complex effects on economic stability [1-3]. Central banks and governments strive to maintain sustainable growth, low unemployment, controlled inflation, and optimal public debt levels, with the strength of the national currency often serving as a direct indicator of economic health. Exchange rate forecasting has long been a significant topic in academic literature. High accuracy and efficiency in this process are of strategic importance for financial markets, national budgets, and international trade.

The application of modern mathematical and statistical methods in the study of economic processes, as well as macro- and microeconomic dependencies, provides valuable insights into currency and financial markets [4,5]. Many of the current approaches combine classical techniques with machine learning to achieve higher predictive accuracy. For example, in [6], the authors propose an ensemble model combining linear and nonlinear techniques (Box–Jenkins ARIMA methods and neural networks) for forecasting major currency pairs (EUR/USD, GBP/USD, USD/JPY). The study shows that the hybrid model is more precise than each of the individual methods. In [7], the predictive power of Random Forests is demonstrated for forecasting four major exchange rates (USD/CAD, GBP/USD, AUD/USD, EUR/USD). Study [8] proposes a new hybrid model of Fuzzy AutoRegressive Integrated Moving Average (FARIMA) integrated with probabilistic neural classifiers (FARIMAH), achieving a more accurate model for financial

forecasting with incomplete data. Empirical results from exchange rate market forecasting indicate that the proposed model effectively improves forecasting accuracy, especially when dealing with scarce data over a short time period. The authors' approach also shows improved performance in the presence of unstable or anomalous data. In [9], a combined model incorporating logistic regression, a random forest classifier, and a Gaussian Naïve Bayes classifier is developed to identify optimal times for buying and selling the euro against the dollar, achieving an accuracy of 94.8%. In [10], the authors propose a new method for predicting USD exchange rates against GBP, JPY, EUR, and CNY by combining Long Short-Term Memory (LSTM) with a bagging ensemble strategy (LSTM-B), aiming to improve both forecasting accuracy and trading profitability. Further studies, such as [11], evaluate the performance of advanced machine learning methods for trading the EUR/USD exchange rate at the ECB fixing. Five supervised classification algorithms (K-Nearest Neighbours, Naïve Bayes, Artificial Neural Networks, Support Vector Machines, and Random Forests) are applied to predict the one-day-ahead movement of the exchange rate using only autoregressive variables. Their performance is compared with two traditional approaches (Naïve Strategy and MACD), showing that SVM and Random Forest yield the highest returns and Sharpe ratios. Paper [12] focuses on the foreign exchange market, known for its high volatility and large transaction volumes, emphasising the importance of accurately predicting currency pair movements for effective investment strategies. It presents a comparative analysis of various machine learning models for forecasting the daily directional movement of the EUR/USD pair, evaluating both decorrelated and non-decorrelated feature sets via Principal Component Analysis. The work in [13] highlights that previous exchange rate forecasting research has typically relied on either shallow or deep features, with limited investigation into their combined potential. To address this, it proposes a novel ensemble learning method, Adaptive Linear Sparse Random Subspace (ALS-RS), which leverages the complementary strengths of both feature types.

Ensemble learning represents a class of machine learning techniques that combine multiple base models to achieve superior predictive performance compared to individual models [14,15]. By aggregating diverse learners – either homogeneous, as in bagging methods (e.g., Random Forest), or heterogeneous, as in stacking and boosting approaches – ensemble methods can reduce variance, bias, or both, depending on the strategy employed. This approach is particularly effective in financial time series forecasting, where high volatility and noise can lead to instability in single-model predictions. In the context of exchange rate forecasting, ensemble learning makes it possible to integrate models that capture different aspects of the data (such as linear and nonlinear relationships and short- and long-term dependencies), thereby enhancing overall prediction accuracy and robustness.

The primary objective of this study is to develop high-performance machine learning models for predicting the USD/JPY exchange rate, while examining the impact of key macroeconomic indicators for both countries. For this purpose, time series analysis is combined with CART Ensemble and Arcing methods.

2. Data and methods

2.1. Data collection

In this study, we use official data provided by the Bloomberg Professional Terminal. The dataset is arranged by month and spans the period from October 1987 to April 2025. The USD/JPY exchange rate (*EX_USDJPY*) is treated as a dependent variable, while a set of macroeconomic indicators serves as predictors. Table 1 summarises the variables used, along with their descriptive statistics. An important characteristic of the data in Table 1 is that all time series are measured simultaneously, which makes real-time forecasting particularly challenging.

Table 1. Descriptive statistics of the data used[1].

Variable	Description	Measure	Mean	Median	Std. deviation
FX_USDJPY	US dollars to Japanese yen Exchange Rate	Rel. units	114.69	112.69	17.57
IR_US	Fed Interest Rate	%	3.23	3.00	2.58
IR_JP	BOJ Interest Rate	%	0.79	0.15	1.47
INF_US	US Inflation Rate	%	2.81	2.7	1.62
INF_JP	Japan Inflation Rate	%	0.72	0.3	1.37
UNEM_US	US Unemployment Rate	%	5.69	5.4	1.69
UNEMJ_P	Japan Unemployment Rate	%	3.52	3.4	1.03
BOND_US	US 10-Year Government Bond Yield	%	4.49	4.25	2.18
BOND_JP	Japan 10-Year Government Bond Yield	%	1.92	1.36	1.91

[1]All percentages are measured on a monthly basis.

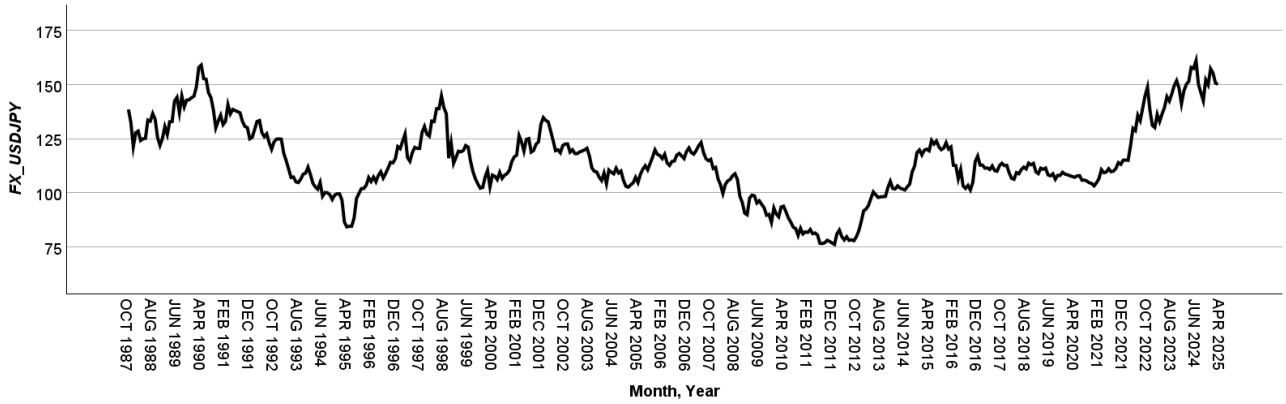


Figure 1. USD/JPY exchange rate from October 1987 to April 2025.

Figure 1 displays the time series of the dependent variable FX_USDJPY, from which three distinct phases of the Japanese yen can be observed:

(1) a period of yen appreciation (1990 – 2011), during which the exchange rate steadily declined to around 75–80 JPY per USD;
(2) a period of significant depreciation (2012 – 2020), when the rate rose sharply, exceeding 125 JPY per USD around 2015; and
(3) a post-2022 period marked by a clear trend of yen weakening.

2.2. Methods and Model Evaluation Measures

The modelling framework is carried out through a two-stage procedure consisting of feature selection followed by model specification. In the first stage, a time-series analysis is conducted on the differential dynamics of the macroeconomic indicators for the two economies, using standard approaches such as the parametric univariate Box–Jenkins ARIMA methodology [15] and partial autocorrelation function (PACF) analysis. Once the relevant predictors (features) are determined for modelling and forecasting *FX_USD/JPY*, we apply the powerful ensemble machine learning techniques of CART-Ensembles and Arcing. The CART Arcing method (Adaptive Resampling and Combining), commonly known as Arcing, is an ensemble approach in which multiple decision trees (CART – Classification and Regression Trees) are generated and combined to improve predictive accuracy in both classification and regression tasks. Arcing belongs to the class of boosting algorithms, of which the best-known example is Arc-x4, proposed in [16]. While traditional ensemble methods with Bagging assign equal probability to each observation when creating bootstrap samples, Arcing adjusts these probabilities dynamically, based on how frequently each observation has been misclassified in previous trees. This allows the model to focus more heavily on complex cases. Each successive tree is thus trained in a way that improves upon the performance of the preceding ones. The weight assigned to each observation is updated according to the formula: $w_{t+1}(V_i) = \frac{1+m(V_i)^4}{\sum_i(1+m(V_i)^4)}$, where $m(V_i)$ denotes the number of times observation V_i was misclassified by the models built up to iteration $1, 2, \ldots, t$.

The primary hyperparameters of the Arcing ensemble include: the number of trees in the ensemble (T), form of cross-validation (CV), the minimum admissible size of parent nodes ($m1$) and child nodes ($m2$), and the number of predictors randomly selected at each node split. The principal advantages of Arcing-based ensembles lie in their higher predictive accuracy, resistance to overfitting, and robustness to noise.

To evaluate model performance, we use the following statistical measures: Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination R^2:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2}{n}}, \quad (1)$$

$$MAPE = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{P_k - Y_k}{Y_k}\right|, \quad (2)$$

$$R^2 = \frac{\{\sum_{k=1}^{n}(P_k - \bar{P})(Y_k - \bar{Y})\}^2}{\sum_{k=1}^{n}(P_k - \bar{P})^2 \sum_{k=1}^{n}(Y_k - \bar{Y})^2}, \quad (3)$$

where Y_k and P_k are the values of the dependent and predictor variables, respectively, and $\bar{Y}$ and $\bar{P}$ are their means. Finally, n is the sample size.

3. Modelling Results

3.1. Feature Selection

Differences (differentials) between macroeconomic indicators of two countries reflect their relative economic conditions and form the basis

of many classical and modern models of exchange rate determination. According to the Interest Rate Parity (IRP) theory, the expected change in the exchange rate between two countries is determined by the difference in their nominal interest rates, i.e. the interest rate differential. A higher interest rate in one country tends to attract capital inflows, leading to appreciation of the domestic currency [17]. The theory of Purchasing Power Parity (PPP) predicts that the exchange rate should adjust in line with inflation differences between countries (the inflation differential): a country with higher inflation will experience depreciation relative to a lower-inflation partner [18]. Although not directly related to the exchange rate, the unemployment differential between two countries serves as an indicator of economic activity, as lower unemployment often implies a prospect of higher interest rates in the future, which in turn affects the currency [19]. Finally, the bond yield differential reflects both interest rate risk and expectations of economic stability and inflation: higher yields tend to attract foreign capital and strengthen the currency [20].

In the present study, to examine the influence of macroeconomic indicators on the USD/JPY exchange rate, the following variables were defined:

- *IR_DIFF = IR_US – IR_JP* (Interest Rate Differential)
- *INF_DIFF = INF_US – INF_JP* (Inflation Differential)
- *UNEMP_DIFF = UNEMP_US – UNEMP_JP* (Unemployment Differential)
- *BOND_DIFF = BOND_US – BOND_JP* (Bond Yield Differential)

Since exchange rate movements are jointly determined with macroeconomic indicators, selecting appropriate predictors is essential for constructing an effective forecasting exchange rate model.

Figure 2 presents the PACFs of the differenced macroeconomic indicators for the USA and Japan. As observed from the graphs, all four variables under consideration (*IR_DIFF*, *INF_DIFF*, *UNEMP_DIFF*, *BOND_DIFF*) exhibit a pronounced first lag, indicating a very strong and statistically significant partial autocorrelation (close to 1), i.e., a dependency between the current time series and the same series shifted by one lag. For the variable *IR_DIFF*, lags 2, 3, and 4 are also significant, and for *INF_DIFF*, lags 2 and 13 are significant. All of these are potential predictors for constructing a model of the USD/JPY exchange rate. We denote the lagged variables as *LIR_DIFF*, *LINF_DIFF*, *LUNEMP_DIFF*, and *LBOND_DIFF* for the first lag, and as *Lk_<var>* for the k-th lag of the corresponding variable.

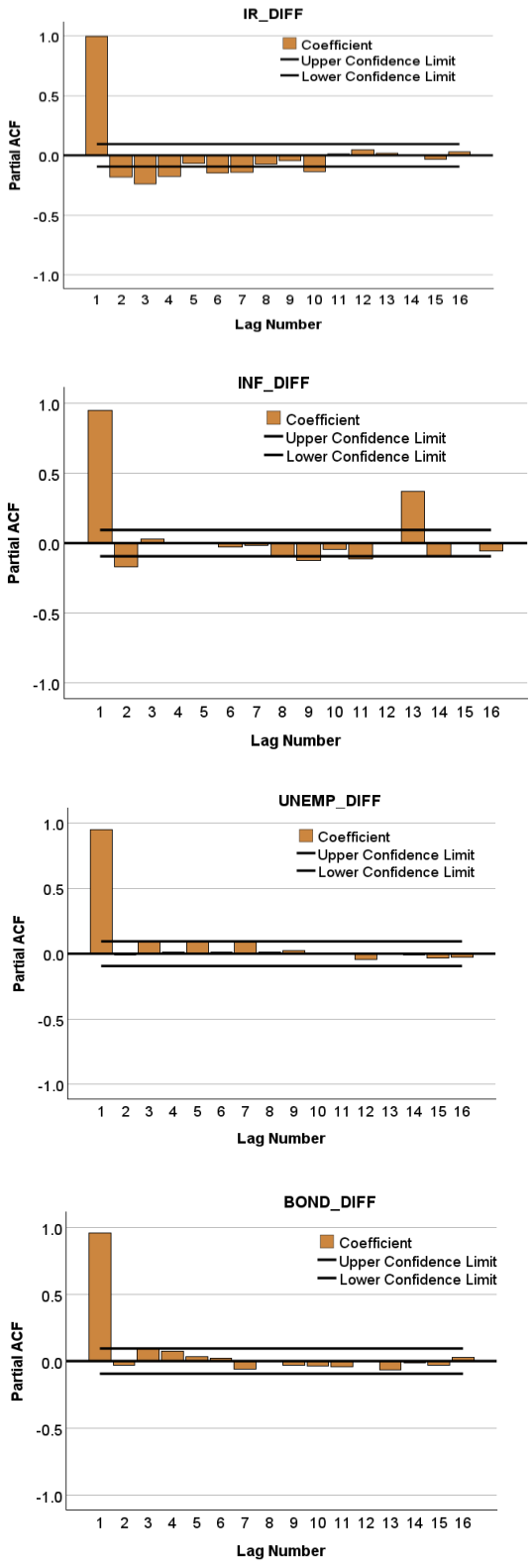


Figure 2. PACF of the differentials and macroeconomic indicators of the USA and Japan.

Figure 3 presents the PACF of the dependent variable *FX_USDJPY*. The strongest partial autocorrelation appears at the first lag, exceeding the boundaries of the confidence interval, which indicates that the previous value of the exchange rate (*LFX_USDJPY*) is a significant predictor of its current

value. The other lags are close to zero and lie within the confidence limits, i.e., they are not statistically significant. This suggests that there is no substantial dependence between earlier values of the variable and its current value.

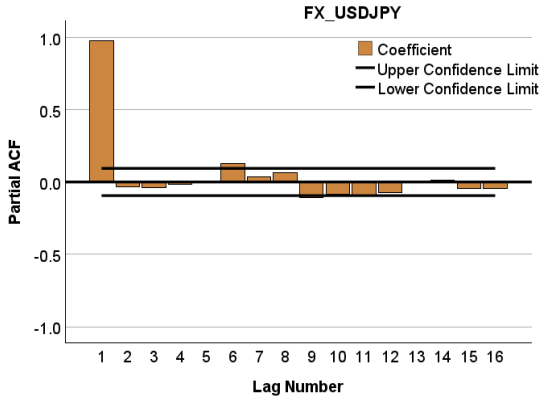


Figure 3. PACF of the USD/JPY exchange rate.

For completeness of this study, Table 2 reports the univariate SARIMA *(p, d, q) (P, D, Q)* models of the main variables. A *SARIMA* model consists of two main components: a non-seasonal part *(p, d, q)* and a seasonal part *(P, D, Q)*, which together capture both short-term dynamics and seasonal patterns in the time series data. The non-seasonal components are: *p* (autoregressive order), i.e., the number of lagged observations included in the model (AR part); *d* (degree of differencing), i.e., the number of times the data is differenced to achieve stationarity; and *q* (moving average order), i.e., the number of lagged forecast errors in the prediction equation (MA part). The seasonal components *(P, D, Q)* are, respectively: *P* (seasonal autoregressive order), i.e., the number of seasonal lagged observations (e.g., 12 months ago); *D* (seasonal differencing), i.e., the number of seasonal differences required to remove seasonal trends; and *Q* (seasonal moving average order), i.e., the number of seasonal lagged forecast errors.

Table 2. Statistics of the obtained univariate SARIMA models for differential dynamics of the macroeconomic indicators.

Variable	SARIMA $(p,d,q)(P,D,Q)_s$	Ljung-Box statistic	R^2
FX_USDJPY	(0, 1, 5) $(0,0,0)_{12}$	0.288	0.957
IR_DIFF	(0, 1, 9) $(0,0,0)_{12}$	0.568	0.990
INF_DIFF	(0, 1, 1) $(0, 0, 1)_{12}$	0.580	0.942
UNEMP_DIFF	(1, 1, 2) $(0,0,0)_{12}$	0.986	0.904
BOND_DIFF	(0, 1, 0) $(1, 0, 1)_{12}$	0.665	0.918

The Ljung-Box (L-B) statistics of all models are insignificant; therefore, we can assume the models are adequate [15]. Additionally, the models exhibit excellent performance with a high R^2 accuracy rate. The key takeaway from Table 2 is that all variables exhibit a differencing order of $d = 1$, indicating that they are non-stationary in levels but become stationary after their first differencing. This property supports the use of lagged versions of these variables as meaningful predictors in a future model construction.

3.2. Construction and Evaluation of CART-Ensembles and Arcing Models

We model the USD/JPY exchange rate across several time periods, denoted as: Period I (January 1987 – April 2025), Period II (January 2000 – April 2025), Period III (January 2012 – April 2025), and Period IV (January 2020 – April 2025). Period I encompasses all available data on the exchange rate and macroeconomic indicators, totalling $n1 = 451$ cases. Period II captures the exchange rate dynamics in the third millennium, with $n2 = 304$ cases. Period III corresponds to the period when the weakening of the yen began, with $n3 = 160$ cases. Period IV covers the post-COVID period, marked by a clear trend of yen depreciation, with $n4 = 64$ cases. To demonstrate the robustness of the obtained models, for each period, we remove the last four monthly USD/JPY observations and predict them step by step, one month ahead, using the selected model.

For each period, multiple models were constructed using different selections of hyperparameters for the Arcing ensemble and various groups of lagged variables. Based on statistical measures, predictive performance, predictor significance, and residual analysis, we selected the following set of lagged variables: *LFX_USDJPY*, *LIR_DIFF*, *L3_IR_DIFF*, *LINF_DIFF*, *LUNEMP_DIFF*, *LBOND_DIFF*. The models were built using the following hyperparameter settings: number of trees in the ensemble $T = 30, 40, 50, 60, 70$; type of cross-validation (CV) – CV5 and CV10; minimum observations in parent and child nodes of each tree of each ensemble $m1{:}\,m2 = 5{:}\,5$; and number of randomly selected predictors for building each tree – 3. The Arcing models that achieved the best statistical performance and predictive accuracy on the test set were obtained for all four periods using $T = 60$ trees, CV10 cross-validation, and minimum cases in parent and child nodes $m1{:}\,m2 = 5{:}\,5$.

For each of the four periods, we constructed four models to predict the USD/JPY exchange rate – separately for January, February, March, and April 2025. Since the four models for a given period exhibited similar statistical performance, Table 3 presents the metrics for the weakest model, which we denote as Mi, where i corresponds to the period number. The other models for the same period achieved better performance, as indicated by the symbols "<" and ">".

Table 3. Statistics of the CART-Ensembles and Arcing models.

Model	RMSE	MAPE	R^2
M1	< 2.593	< 0.019	> 0.978
M2	< 1.328	< 0.011	> 0.994
M3	< 1.382	< 0.011	> 0.995
M4	< 1.684	< 0.011	> 0.992

The results in Table 3 indicate excellent model performance, with a coefficient of determination up to 98% and MAPE of 1%.

Table 4 presents the step-by-step forecasts of the models for the four periods, spanning January 2025 to April 2025.

Table 4. Predicted values of the models for the period January 2025 – April 2025.

Predicted	M1	M2	M3	M4	Real value
January 2025	153.52	153.36	152.72	153.67	155.19
February 2025	152.83	152.71	151.89	152.37	150.63
March 2025	151.78	151.88	151.72	152.03	149.96
April 2025	150.93	150.74	150.23	151.05	143.07

The predicted values for models *M1*, *M2*, *M3*, and *M4* fit the actual data with approximately 88%, 93%, 99%, and 95% respectively.

Figure 4 illustrates the sequence plot of the original USD/JPY data together with the values predicted by model *M3*. Following the vertical line are the step-by-step forecasts produced by model *M3* one month ahead, applied to the months of January, February, March, and April 2025, respectively.

For greater clarity, Figure 5 displays the actual data and the values predicted by model *M3* over the last 16 months, with the step-by-step forecasts again marked by a vertical line.

Figure 6 shows the quality of fit between the original USD/JPY data and the predictions of model *M3*, along with the 95% confidence intervals. An excellent approximation is observed, including the extreme values.

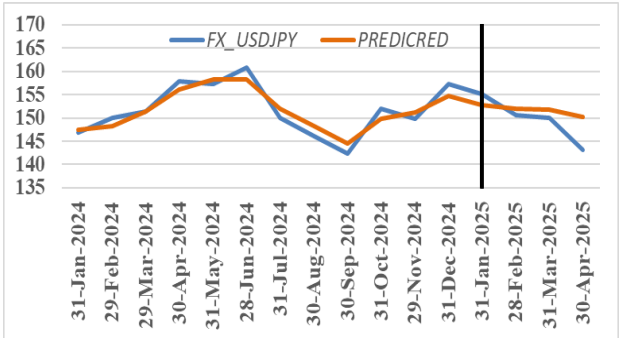


Figure 5. Original USD/JPY data and forecasts from model M3 for the period January 2024 – April 2025.

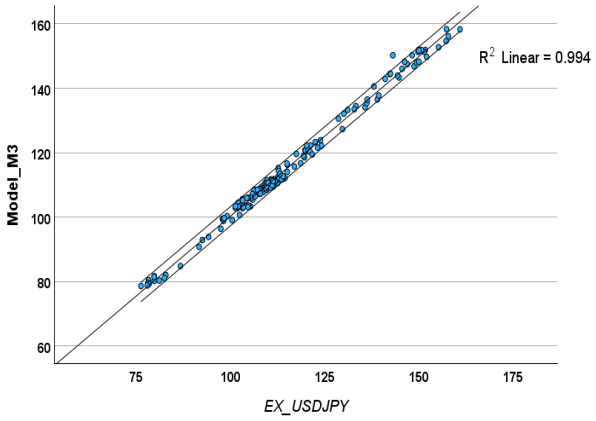


Figure 6. Scatter plot of the USD/JPY data and predictions from the M3 Arcing model.

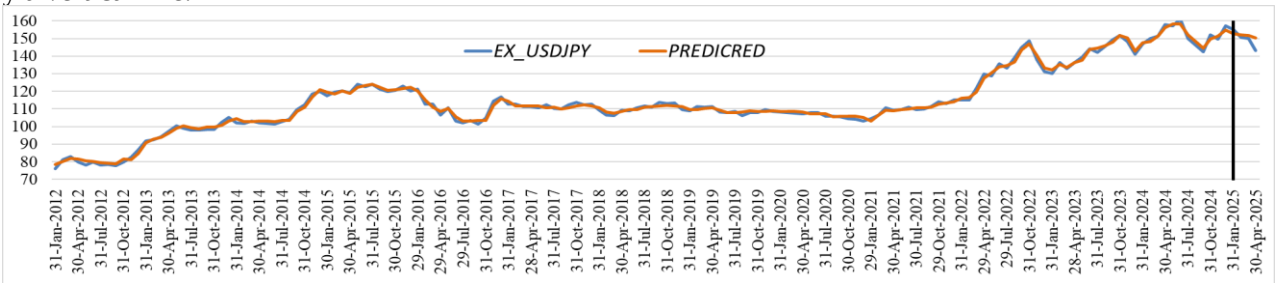


Figure 4. Original USD/JPY data and forecasts from model M3.

3.3. Variable importance

The CART-Ensemble and Arcing method allows us to determine the relative influence of the predictors in the model. For each of the four periods, the models constructed for step-by-step prediction of the last four monthly values determine the same ranking of predictors. The models for the different periods determine the order of predictor influence as shown in Table 5. For all models, the first variable with a weight of 100 points is the lagged dependent variable LFX_USDJPY. The next most important variable for models *M1* and *M2* is the lagged interest rate differential at one and three lags, and both have almost equal influence percentages. With a difference of up to 5%, the bond differential variable follows, then the unemployment differential, and finally the inflation differential. For models *M3* and *M4*, the only difference is that the bond differential precedes, by up to 5%, the lagged interest rate differential variables at one and three lags.

Table 5. Relative variable importance of predictors in the models.

Variable	Relative importance in %			
	M1	M2	M3	M4
LFX_USDJPY	100	100	100	100
LIR_DIFF	35	49	72	80
L3_IR_DIFF	35	45	68	78
LBOND_DIFF	28	42	80	90
LUNEMP_DIFF	25	35	80	65
LINF_DIFF	15	15	20	45

3.4. Reliability Evaluation of the Forecasts

Following [19,20], we need to evaluate the reliability of all forecasting models. To this end, an Auto-Correlation Function (ACF) plot can be used to check for the presence of autocorrelation in the model residuals. We verify whether the residuals are identically distributed, asymptotically independent, and whether there is no serial correlation remaining in the model errors.

The ACF of the residuals of model *M3* in Figure 7 shows no deviations outside the confidence intervals. The residual plots for the other models exhibit similar behaviour. We can therefore conclude that the selected predictors accurately forecast the USD/JPY exchange rate, regardless of the chosen time period.

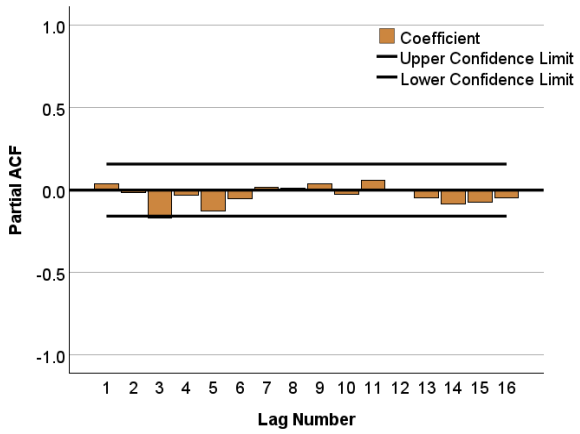


Figure 7. Autocorrelation function of M3 residuals.

From an economic perspective, the results indicate that differentials in interest rates, inflation, and bond yields remain key drivers of USD/JPY movements. The persistent dominance of interest rate and bond yield differentials implies that monetary policy asymmetries between the Federal Reserve and the Bank of Japan are the main determinants of exchange rate fluctuations. The strong predictive contribution of these variables also highlights the sensitivity of the yen to shifts in global capital flows and policy expectations. These findings suggest that ensemble learning models not only achieve high forecasting accuracy but also effectively capture the economic structure underlying exchange rate formation.

4. Conclusion

This study demonstrates the high performance of the CART-Ensembles and Arcing method for forecasting the exchange rate. The main contribution of this study lies in applying the Arcing ensemble method with a carefully selected set of macroeconomic predictors. The results confirm that the method maintains strong statistical properties across varying time periods.

In this article, we established the dependence of changes in the USD/JPY exchange rate on the differentials of various macroeconomic indicators for the two countries. Accurate prediction of future exchange rate values cannot be achieved without considering the key macroeconomic indicators of both economies.

Future research could extend this framework to other major currency pairs (e.g., EUR/USD, GBP/JPY) or combine ensemble learning with deep neural architectures such as LSTM or Transformers to better capture nonlinear temporal dependencies. Incorporating sentiment indices, geopolitical risk indicators, and policy uncertainty measures could also enhance the robustness of exchange rate predictions.

Acknowledgements

This study was supported by the Bulgarian National Science Fund, Grant KP-06-N52/9.

References

[1] Byrne, J. P., Korobilis, D., & Ribeiro, P. J. (2016). Exchange rate predictability in a changing world. *Journal of International Money and Finance, 62*, 1–24. https://doi.org/10.1016/j.jimonfin.2015.12.001

[2] Mishkin, F. S. (2012). *Economics of money, banking, and financial markets* (10th ed.). Prentice Hall.

[3] Parkin, M. (2011). *Economics* (10th ed.). Prentice Hall.

[4] Elliott, G., & Timmermann, A. (Eds.). (2013). *Handbook of economic forecasting* (Vol. 2). North Holland.

[5] Fabozzi, F. J., & Mann, S. V. (2012). *The handbook of fixed income securities.* McGraw-Hill Education.

[6] Tripathi, M., Kumar, S., & Inani, S. K. (2021). Exchange rate forecasting using ensemble modeling for better policy implications. *Journal of Time Series Econometrics, 13*(1), 43–71. https://doi.org/10.1515/jtse-2020-0013

[7] Tsuji, C. (2022). Exchange rate forecasting via a machine learning approach. *iBusiness, 14*(3). https://doi.org/10.4236/ib.2022.143009

[8] Khashei, M., Rafiei, F. M., & Bijari, M. (2013). Hybrid fuzzy auto-regressive integrated moving average (FARIMAH) model for forecasting the foreign exchange markets. *International Journal of Computational Intelligence Systems, 6*, 954–968. https://doi.org/10.1080/18756891.2013.809937

[9] El Mahjouby, M., Taj Bennani, M., Lamrini, M., Bossoufi, B., Alghamdi, T. A. H., & El Far, M. (2024). Machine learning algorithms for forecasting and categorizing euro-to-dollar exchange rates. *IEEE Access, 12*, 74211–74217. https://doi.org/10.1109/ACCESS.2024.3404824

[10] Sun, S., Wang, S., & Wei, Y. (2020). A new ensemble deep learning approach for exchange rates forecasting and trading. *Advanced Engineering Informatics, 46*, 101160. https://doi.org/10.1016/j.aei.2020.101160

[11] Theofilatos, K., Likothanassis, S., & Karathanasopoulos, A. (2012). Modeling and trading the EUR/USD exchange rate using machine learning techniques. *Engineering, Technology & Applied Science Research, 2*(5), 269–272. https://doi.org/10.48084/etasr.200

[12] Guyard, K. C., & Deriaz, M. (2024). Predicting foreign exchange EUR/USD direction using machine learning. In *Proceedings of the 2024 7th International Conference on Machine Learning and Machine Intelligence (MLMI '24)* (pp. 1–9). Association for Computing Machinery. https://doi.org/10.1145/3696271.3696272

[13] Wang, G., Tao, T., Ma, J., Li, H., Fu, H., & Chu, Y. (2021). An improved ensemble learning method for exchange rate forecasting based on the complementary effect of shallow and deep features. *Expert Systems with Applications, 184*, 115569. https://doi.org/10.1016/j.eswa.2021.115569

[14] Zhou, Z.-H. (2012). *Ensemble methods: Foundations and algorithms.* CRC Press. https://doi.org/10.1201/b12207

[15] Opitz, D., & Maclin, R. (1999). Popular ensemble methods: An empirical study. *Journal of Artificial Intelligence Research, 11*, 169–198. https://doi.org/10.1613/jair.614

[16] Breiman, L. (1998). Arcing classifiers. *The Annals of Statistics, 26*(3), 801–849. https://doi.org/10.1214/aos/1024691079

[17] Obstfeld, M., & Rogoff, K. (1996). *Foundations of international macroeconomics.* MIT Press.

[18] Taylor, A. M., & Taylor, M. P. (2004). The purchasing power parity debate. *Journal of Economic Perspectives, 18*(4), 135–158. https://doi.org/10.1257/0895330042632744

[19] Engel, C., & West, K. D. (2005). Exchange rates and fundamentals. *Journal of Political Economy, 113*(3), 485–517. https://doi.org/10.1086/429137

[20] Bernoth, K., & Erdogan, B. (2012). Sovereign bond yield spreads: A time-varying coefficient approach. *Journal of International Money and Finance, 31*(3), 639–656. https://doi.org/10.1016/j.jimonfin.2011.10.006

Computer Science and Interdisciplinary Research Journal
Volume 2, Issue 2, 2025. ISSN: 3033-1218, DOI: 10.70862/CSIR.2025.0202-17

Integration of GenAI and Behavioral Analysis in an Authentication System

Georgi Atanasov, Kostadin Yotov, Emil Hadzhikolev

University of Plovdiv Paisii Hilendarski, 236, Bulgaria Blvd, 4027 Plovdiv, Bulgaria

stu2501322014@uni-plovdiv.bg, kostadin_yotov@uni-plovdiv.bg, hadjikolev@uni-plovdiv.bg

Abstract – This paper presents a model of a security-enhanced web authentication system that combines behavioral analysis, generative artificial intelligence, and an event-driven architecture. The aim of the study is to develop a dynamic monitoring and alerting mechanism capable of identifying unusual activity based on contextual and behavioral indicators, without imposing additional authentication steps on the user. The system leverages Redis with Pub/Sub and TTL notifications for reactive event processing, generative AI for automatic summarization of logs into a human-readable format, and the Isolation Forest algorithm for experimental detection of behavioral anomalies. The results demonstrate that the proposed architecture can effectively identify suspicious sessions, provide interpretable textual explanations of user actions, and react to key events in real time while maintaining low latency.

Keywords – authentication, generative artificial intelligence, GenAI, behavioral security, reactive architecture, anomaly detection.

1. Introduction

Over the past two decades, the internet ecosystem has experienced exponential growth – both in the number of active users and in the diversity and complexity of online services. The digitalization of business, education, and the public sector has led to a situation in which vast amounts of sensitive data – personal, financial, and professional – are transmitted, stored, and processed in networked environments [1-3]. In this context, authentication and authorization have become critical components of cybersecurity, determining not only the technical reliability of a system but also the level of trust placed in it by end users.

Traditional authentication methods, based solely on username–password combinations, can no longer address the dynamic threats of the modern digital landscape. Password leaks, phishing attacks, automated credential-stuffing bots, and the misuse of compromised accounts have become increasingly common [4-7]. According to the Verizon Data Breach Investigations Report [8], more than 70% of system breaches begin with unauthorized access obtained through compromised user credentials. This highlights the vulnerability of systems that rely on static authentication mechanisms.

Among the leading trends in authentication systems are multi-factor authentication (MFA), behavioral biometrics, and dynamic risk assessment, all of which provide continuous user validation [9-11]. Modern web applications no longer rely on a single initial login check; instead, they incorporate ongoing monitoring of user behavior, location, and access context.

An additional challenge stems from the increasing use of artificial intelligence (AI) and automated services acting on behalf of the user, which necessitates systems capable of distinguishing human actions from machine-generated ones. As a result, authentication is evolving from a static entry point into a dynamic trust mechanism that adapts its responses based on context, risk levels, and the history of interactions.

In this technological and social context, there arises a need for a new generation of authentication and access-control systems – ones that combine classical

DOI: https://doi.org/10.70862/CSIR.2025.0202-17

Corresponding author: Emil Hadzhikolev,
University of Plovdiv Paisii Hilendarski
Email: hadjikolev@uni-plovdiv.bg

Received: 30 November 2025.
Revised: 20 December 2025.
Accepted:21 December 2025.
Published: 21 December 2025.

BY-NC-ND © 2025 Georgi Atanasov, Kostadin Yotov, Emil Hadzhikolev. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

The article is published with Open Access at http://csirjournal.com/

security principles with behavioral analysis, automation, and generative artificial intelligence (GenAI). Such systems not only validate the user but also understand the way the user interacts with the application, turning security into an intelligent and adaptive process.

The aim of the present study is to develop a dynamic and adaptive authentication and access-control system that enhances security trust through the integration of three key components:

- Behavioral security – monitoring request frequency, navigation paths, session duration, geographic location, and device type. These characteristics form a "digital behavioral fingerprint" that enables the detection of deviations typically associated with misuse or malicious activity.
- GenAI – automatic summarization of user sessions into clear, contextual reports that accelerate analysis and response.
- Intelligent alerting – automatic notifications triggered by anomalies, risks, or unusual changes in access patterns.

The study proposes a model of a dynamic and adaptive authentication system – Forest Gate. A prototype has been developed, with its source code publicly available at: https://github.com/georgi2005atanasov/forest_gate. Experiments conducted with this prototype show that the system operates reliably and stably in real-time across different user-behavior scenarios.

2. Technologies and Concepts for Authentication and Access Control

Authentication and access-control systems are among the oldest yet fastest-evolving components of modern cybersecurity. From simple username–password authentication to multifactor and context-aware approaches, developments in this field reflect the continuous pursuit of three key objectives: reliability, usability, and adaptability.

2.1. Classical and Modern Authentication Mechanisms

The classical authentication model based on a username and password has long been the dominant standard in web security [12-14]. However, its simplicity and ease of integration come with a significant drawback – the absence of adaptability to an evolving threat landscape. In an environment characterized by a growing number of cyberattacks, automated password-cracking tools, and social engineering techniques, static authentication methods are becoming increasingly ineffective.

A password represents a single factor of trust – knowledge that can be compromised through phishing, database leaks, or brute-force attacks. According to IBM Security [15], the average time to detect a breach involving stolen credentials exceeds 200 days, allowing attackers to operate unnoticed for extended periods.

Classical login mechanisms also suffer from a lack of context. The system does not distinguish whether the user is logging in from a familiar device, from an unusual geolocation, or with atypical behavioral patterns – for example, rapid navigation, high-volume API requests, or activity during abnormal hours. Therefore, even a successfully completed authentication does not guarantee that the access is legitimate [16-18].

Modern attacks such as credential stuffing, session hijacking, token replay, and keylogging exploit precisely this static nature. Since traditional systems do not track contextual factors, they cannot detect cases in which a valid token is used simultaneously from different countries – a scenario commonly referred to as impossible travel. In such situations, the absence of behavioral monitoring and dynamic risk assessment results in a delayed or entirely ineffective system response.

With the evolution of internet applications and web services requiring shared access across multiple clients and servers, new standards for secure identification have emerged. Among the most prominent are OAuth 2.0 and OpenID Connect, which provide mechanisms for delegated access through access tokens [19-21]. These standards allow users to authenticate themselves to different applications without revealing their passwords and enable centralized rights management. JSON Web Tokens (JWTs) have become the de facto standard for portable, self-contained tokens in distributed architectures and microservice environments [19,22].

Although these standards address important security challenges, they remain static in nature: authentication is performed only once, while access remains valid until the token expires. This makes them vulnerable to token theft, session reuse, and session hijacking attacks.

2.2. Behavioral Analysis and Contextual Authentication

Among contemporary trends, the use of behavioral analytics as a complement to classical authentication mechanisms stands out. This approach considers not only who is accessing the system but also how they do it – by analyzing the speed and sequence of actions, timing patterns, location, device type, navigation habits, and even typing behavior [9,10,23].

Through statistical models and machine-learning algorithms, the system can detect deviations from typical behavior and indicate potential risk.

Approaches such as continuous authentication and adaptive risk-based authentication enable real-time decision-making – for example, requesting an additional authentication factor or temporarily restricting access when suspicious activity is detected [11,24].

2.3. Use of AI in Security

In recent years, AI has become a key tool for analysis and automation in the security domain. Traditional machine-learning models – such as Random Forest, Support Vector Machines, and Neural Networks – are used for anomaly classification, phishing and spam detection, risk assessment, and attack prediction [25-27].

With the rise of GenAI, new possibilities have emerged – not only for threat detection but also for explainability and data interpretation. Generative models can transform technical logs and raw data into comprehensible textual descriptions that help administrators and researchers gain a faster overview of events [28-30]. This capability is particularly valuable when processing large volumes of data, where human analysis would be too slow or incomplete.

2.4. Anomaly Detection Algorithms: Isolation Forest

Anomaly detection algorithms are a core component of modern security systems. Among them, Isolation Forest (iForest) stands out for its efficiency on large datasets and its low computational complexity. Its principle is based on the idea that anomalies are "few and easier to isolate" – they require fewer partitions during the recursive data-splitting process [31-34].

2.5. Integrated Trust Systems

Recent research in the security domain increasingly focuses on integrated trust-driven architectures [35,36]. These systems combine technical, behavioral, and contextual data to construct a dynamic trust index for each user and each action. This concept forms the foundation of Zero Trust Security – a model in which access is granted not based on static permissions but according to the current risk assessment.

3. Development of Forest Gate

The Forest Gate system demonstrates an integrated security approach in which classical authentication mechanisms are enhanced through behavioral analysis and GenAI. The architecture is modular and distributed – each component is autonomous but communicates with the others through well-defined interfaces, ensuring transparency in data processing and flexibility for future extensions.

The architectural design is guided by three core requirements:

- Access and data security – protected communication between components using cryptographic keys and strictly limited information sharing.
- Traceability and explainability – detailed logs, session records, and summarized action reports that enable both technical and analytical inspection.
- Automation and adaptability – automatic collection of interaction data, AI-driven analysis via GenAI, and alerting in the presence of anomalies or risk events.

The technologies used in the implementation include Rust, PostgreSQL, Redis, NGINX, Docker, MaxMind GeoIP, SendGrid, and GenAI via OpenRouter. The anomaly detection algorithm employed is Isolation Forest.

3.1. Technologies Used

The implementation of the Forest Gate prototype is based on the integration of several complementary technologies that ensure high performance, stability, and system extensibility:

1. *Programming language and core logic – Rust.*

The system's core is implemented in Rust – a language that has become a standard for secure and efficient software development. Its ownership & borrowing model provides memory safety without a garbage collector and eliminates entire classes of vulnerabilities related to improper resource handling. Rust enables reliable implementation of asynchronous processes – request handling, session management, and Pub/Sub communication – without compromising security, which is critical when managing sensitive authentication data.

2. *Data management – PostgreSQL and Redis.*

PostgreSQL serves as the primary relational store for persistent data such as users, devices, login attempts, and configuration settings. Its ACID compliance and robustness make it suitable for critical systems. Redis functions as an in-memory cache and an inter-process communication mechanism. Thanks to Pub/Sub and keyspace notifications, the system can monitor user activity in real-time, react to session expiration, and trigger automated processes for analysis and alerting.

3. *Containerization and infrastructure – Docker and NGINX.*

The prototype is built using Docker Compose, which integrates the individual components into an isolated and reproducible environment. This greatly facilitates local testing and provides a foundation for

further deployment. NGINX operates as a reverse proxy responsible for request routing, caching, and TLS protection.

4. *GenAI – OpenRouter.*

GenAI models integrated via the OpenRouter API are used to analyze and summarize user behavior. They transform raw logs into concise and readable Markdown descriptions, improving the system's transparency and explainability. The prompts are designed to ensure that the model does not fabricate information but extracts only relevant facts – an essential property for internal auditing and security.

5. *Anomaly detection – Isolation Forest.*

Isolation Forest is used experimentally to identify atypical patterns in user behavior. The algorithm enables the detection of suspicious logins and anomalies in request frequency. Although not a central component, it demonstrates the potential to extend Forest Gate toward adaptive risk assessment, where the system responds based on the statistical abnormality of actions.

6. *Intelligent alerting – SendGrid.*

Integration with SendGrid enables automatic notifications triggered by anomalies, configuration changes, or other risk events. This functionality implements the principle of real-time monitoring – the system not only collects data but also actively communicates with administrators.

7. *Geographical data – MaxMind GeoIP.*

The system uses local GeoLite2-City, GeoLite2-ASN, and GeoLite2-Country databases to enrich requests with geographic information – country, city, and internet service provider (ASN/ISP). This enables the detection of scenarios such as impossible travel and sudden changes in network context.

3.2. Architecture

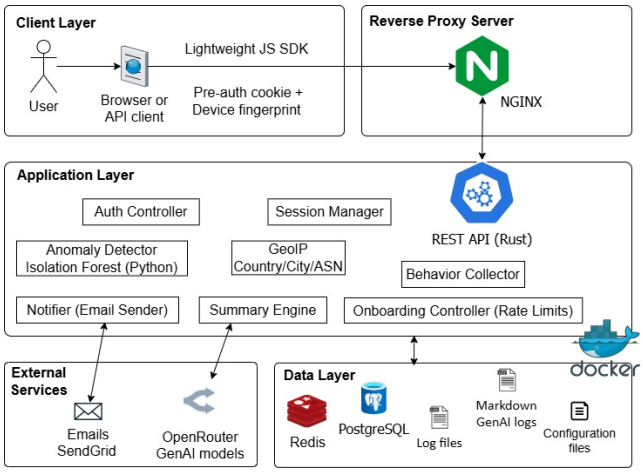


Figure 1. Architecture of Forest Gate.

The architecture of Forest Gate is multilayered and integrates several modules built on the technologies discussed earlier (Figure 1). The components communicate through well-defined interfaces and protocols, ensuring secure, transparent, and adaptive processing of user sessions. The system integrates classical authentication, behavioral analysis, intelligent alerting, and GenAI.

The client layer provides the interface between the user and the system and includes three main components:

- Web application or API client – used to initiate access requests (login, registration, profile updates, etc.).
- Pre-identification system – generates a temporary pre-auth cookie and a unique device fingerprint during the user's first visit. These markers allow the server to associate requests with a specific user and device without compromising anonymity.
- Lightweight JavaScript library – collects behavioral indicators (input timing, navigation speed, request frequency) and forwards them to the backend for analysis.

An NGINX reverse proxy server is positioned as an intermediate layer between the client and the business logic, responsible for request routing, caching, and TLS protection.

The business logic of the system is implemented in the application layer. The core components are developed in Rust, while the Isolation Forest algorithm is implemented in Python. The application server is responsible for authentication, data management and storage, event processing, and communication with the other modules in the architecture. The main modules in the server layer are:

- Onboarding Controller – used during initial registration or a user's first visit. It stores the device fingerprint and activates rate limits based on three parameters: IP address, visitorId (generated on first visit), and the pre-auth cookie.
- Auth Controller – validates credentials and issues sessions/tokens. Sessions can be cached in Redis and persisted when necessary.
- Behavior Collector – receives client-side events and stores them in Redis with a limited TTL. When a key expires, Redis triggers an expired event via Pub/Sub, which initiates subsequent processing.
- Summary Engine (GenAI) – a background service (flusher) that listens for Pub/Sub expiration notifications, collects accumulated events, and generates a concise Markdown summary of the behavior using OpenRouter.
- Anomaly/Risk Module – performs basic risk assessment and detects atypical access patterns using the Isolation Forest algorithm.

- GeoIP Enrichment – enriches login attempts and sensitive actions with country, city, and ASN/ISP data using MaxMind GeoIP.
- Notifier – sends email notifications via SendGrid in cases of critical events such as configuration changes.

The application-layer modules rely on several types of storage located in the data layer:

- PostgreSQL – the primary relational database for persistent data (users, devices, login attempts).
- Redis – a cache and buffer for client-side events, as well as a Pub/Sub mechanism for key-expiration events.
- Markdown logs – human-readable summaries of user interactions generated by GenAI.
- Configuration files and keys – settings for external services (OpenRouter, SendGrid, MaxMind GeoIP) and elliptic-curve cryptographic keys provided through environment variables.

The architecture also includes two external services:

- OpenRouter – used to generate GenAI behavior summaries.
- SendGrid – used to send email notifications for anomalies and configuration changes.

3.3. Data Flow and Component Interaction

The information flow in Forest Gate can be summarized in the following steps:

1. *Session initialization.*

The user accesses the system through the client interface, after which the system generates a pre-auth cookie and a device fingerprint. These enable early device identification and the application of baseline rate limits.

2. *Authentication.*

The Auth Controller validates the provided credentials and creates a session, which is stored in Redis and PostgreSQL. Additionally, the accumulated contextual information (GeoIP data, fingerprint, timestamp, IP address) is passed to Isolation Forest, which determines whether the login attempt is normal or requires blocking or an additional authentication factor (MFA).

3. *Behavior monitoring.*

The Behavior Collector gathers contextual and temporal indicators related to user actions and stores them in Redis with a limited TTL.

4. *Anomaly detection.*

The Anomaly Detector analyzes the collected events using Isolation Forest and computes a real-time risk score.

5. *Summarization and interpretation.*

The Summary Engine listens for TTL expiration events from Redis (Pub/Sub), retrieves the accumulated logs, and sends them to GenAI via OpenRouter. The generative model produces a concise and readable summary of the behavior.

6. *Notification.*

When necessary, the Notifier Module triggers SendGrid and sends an alert to the administrator containing information about the behavior, the detected anomaly, and the recommended action.

7. *Visualization and follow-up actions.*

The data are stored and can be analyzed to detect long-term trends, recurring anomalies, or emerging attack patterns.

3.4. GenAI

An innovative characteristic of Forest Gate is the integration of GenAI as an active component within the security architecture. The goal is not to replace traditional mechanisms for analysis or decision-making, but to add a cognitive layer capable of interpreting and presenting complex system processes in a human-readable form.

In classical authentication models, logs and system events remain in a raw, technical format that is difficult to analyze by non-experts. As the volume and complexity of such data increase, there is a growing need for an intelligent module capable of extracting meaning, context, and trends from fragmented information. This is where GenAI becomes particularly relevant.

Within Forest Gate, GenAI operates as an explanatory and analytical component performing three main functions:

- Support during code development – used to generate, verify, and optimize parts of the Rust program logic through interactive communication with the developer.
- Log and event summarization – automatically transforms system records into concise, structured textual explanations suitable for auditing or for downstream training of security models.
- Support for automated alerting – formulates the content of administrator notifications in natural language, clearly distinguishing critical from non-critical events.

3.4.1. Code Generation, Log Summarization, and Automated Notifications

During the development of the system, GenAI was used to support various aspects of programming:

- Template code generation – modeling REST API structures, handling JSON requests, and implementing middleware layers in Rust.
- Refactoring and optimization – providing suggestions for optimizing functions, reducing

latency, and improving concurrency management through async/await.

- Security verification – automatically identifying potential vulnerabilities in the authentication logic, including incorrect token validation, uninitialized variables, or inconsistent access to shared resources.
- Documentation generation – automatically producing descriptions of functions, structures, and API routes in Markdown format, ensuring consistency between the codebase and the documentation.

One of the most important roles of GenAI in Forest Gate is the automatic summarization of system logs and session data. In traditional systems, administrators must manually review hundreds of log lines – a process that is often slow, labor-intensive, and prone to error. In Forest Gate, GenAI acts as a behavioral interpreter, analyzing the sequence of events within a session and transforming them into a concise and comprehensible textual summary. For example:

The user performed advanced search actions and filtered products by price and rating. They added multiple items to the cart and initiated checkout, then changed the shipping address. No anomalies detected. Risk level: Low.

A burst of rapid actions occurred, including repeated add-to-cart and remove actions and multiple coupon attempts. The user refreshed pages several times and toggled payment options. This is unusual volume but not inherently malicious. Risk level: Medium.

The user attempted multiple security bypasses. They tried CSRF attacks twice, followed by an injection attempt. They also attempted to upload a malicious file and routed their request through a proxy. Risk level: High.

These summaries are generated automatically through system and user instructions defined in the Rust Summary Engine module. The raw data from Redis and PostgreSQL are formatted into contextual JSON and sent to the OpenRouter API, whose response is stored in Markdown format. The resulting summaries are saved as log files and can be used both by administrators and for further analysis or model training.

The third functional role of GenAI relates to intelligent alerting and incident communication. When the Anomaly Detector identifies a deviation from normal behavior (e.g., multiple login attempts from different countries within a short interval), the system sends structured information to the Notifier module, which then triggers two processes:

1. Sending a concise technical alert via the SendGrid API containing the specific parameters of the deviation;
2. Generating a textual summary through GenAI that describes the situation in natural language, for example:

Atypical behavior detected for user g.atanasov: two access attempts from IP addresses in different countries within an 8-minute interval. Likelihood of account compromise: medium. Temporary token suspension is recommended.

3.4.2. System and User Prompts

The interaction between the system and GenAI is managed through two main types of instructions – system prompts and user prompts, which together define the model's role, objectives, and the format of the generated response.

The system prompt defines the model's primary role and its operational constraints. In Forest Gate, it is designed so that the model functions as an intelligent security analysis assistant, with restrictions aimed at producing concise and accurate summaries without hallucinations or fabricated information:

You are a security assistant integrated into the Forest Gate system.
Your role is to analyze user session logs and behavioral data.
Summarize the events objectively, in Markdown format, and avoid any assumptions.
Never include speculative statements or personal interpretations.
Be concise (3–5 sentences maximum) and focus on session context, anomalies, and risk level.

User prompts are generated dynamically during system operation. They contain an interaction identifier and a list of events, and are formatted in JSON. Examples of user prompts include:

```
{
  "interaction_id": "f8d2a13c-18ca-4782-88b6-a8bbe8f309de",
  "event": ["clicked add to cart button", "tried uploading bad file"]
}
{
  "interaction_id": "a1b2c3d4-e5f6-7890-abcd-1234567890aa",
  "event": ["visited /shop", "clicked add to cart on Soap", "tried CSRF"]
}
{
  "interaction_id": "e5f6a7b8-c9d0-1234-ef01-5678901234ee",
  "event": ["visited /login", "attempted login username=user1", "attempted login username=user1", "rate limited /login"]
}
```

4. Experiments

The conducted experiments aimed to evaluate the extent to which the Forest Gate prototype can operate stably, reliably, and in real-time under different user-behavior scenarios. The tests were performed in a controlled environment using Docker, Redis,

PostgreSQL, and NGINX, while Postman, Swagger, and LucidChart were used to simulate realistic workloads and visualize the processes.

4.1. Experimental Environment and Methodology

The system was deployed using Docker Compose, with each service (application, database, cache, and reverse proxy) running in its own isolated container.

Technologies and tools used:

- PostgreSQL – primary storage for users, sessions, and login attempts;
- Redis – event caching and implementation of the Pub/Sub mechanism for activity monitoring;
- NGINX – reverse proxy simulating production-like conditions;
- SendGrid API – for sending notifications upon configuration changes;
- MaxMind GeoIP – for geolocation enrichment and detection of "impossible travel" cases;
- OpenRouter (GenAI) – for summarizing sessions and generating textual reports;
- Postman and Swagger – for functional testing and API endpoint validation.

The simulations included several access scenarios:

- standard login and logout from a familiar device;
- multiple unsuccessful login attempts;
- access attempts from different geographical locations within a short time interval;
- high-frequency event generation by manipulating TTL keys in Redis.

4.2. Performance and Reliability

Experiments were conducted to evaluate the behavior of the Redis Pub/Sub mechanism, the efficiency of GeoIP analysis, and the system's performance under high load.

4.2.1. Redis Pub/Sub Operation

During stress tests involving numerous concurrent sessions, Redis demonstrated stable performance, with expired events being processed in under 50 ms. Excellent synchronization was observed between the background flusher process and the Markdown generator – upon TTL expiration, the system automatically extracted all recorded actions, removed them from the cache, and generated a concise summary using GenAI. No event retention issues or missed notifications were detected.

4.2.2. GeoIP Analysis Efficiency

The geolocation enrichment module (MaxMind) correctly identified locations with over 98% accuracy. "Impossible travel" scenarios (e.g., sequential logins from Sofia and Tokyo within 10 minutes) were successfully detected and reflected in the logs. The GenAI-generated summaries clearly highlighted such cases, for example:

> *Detected login from a distant location inconsistent with previous session. Possible credential sharing or VPN usage.*

4.2.3. High-Load Performance

The Rust backend maintained low latency under concurrent requests (35–40 ms per request with 100 simultaneous users). Latency was measured using the formula:

$$L_{avr} = \frac{1}{N}\sum_{i=1}^{N}(t_{out,i} - t_{in,i}), \quad (1)$$

where L_{avr} is the average latency (ms), N is the number of processed requests, and $t_{out,i}$ and $t_{in,i}$ are the completion and arrival times of the i-th request, respectively.

The use of async/await and Rust's memory safety mechanisms made it possible to avoid memory leaks and race conditions that are common in other languages. In simulations with 100 concurrent users, the calculated value was $L_{avr} \approx 37ms$, confirming the system's low latency and the efficiency of Rust's asynchronous request handling.

To measure throughput, the following formula was used:

$$T = \frac{N}{\Delta t}, \quad (2)$$

where T is the throughput (number of requests per second), N is the total number of processed requests, and Δt is the measurement time interval (in seconds).

In the performed experiments, when the /users/login endpoint processed 1,303 requests in 317.1 seconds, the measured throughput was $T = 4.11\,req/s$, which corresponds to service behavior with long-running requests.

To express the system's reliability under load, the classical definition of the success coefficient was used:

$$R = 1 - \frac{E_f}{E_t}, \quad (3)$$

where $R \in [0,1]$ is the system reliability, E_f is the number of failed or interrupted requests, and E_t is the total number of requests.

During testing of /users/login with 1,303 requests and 0 errors, the reliability was $R = 1.0000$, which demonstrates stable system behavior under load.

4.3. Summarizing Results with GenAI

The most interesting part of the experiments was related to the behavior of GenAI when processing logs. For each completed session, Redis forwarded

structured data to the Summary Engine module, which then sent a request to the OpenRouter API. In this context, GenAI can be viewed as a data-transformation function – from structured logs to a semantic summary. According to the classical formulation, this transformation can be expressed as:

$$S = \Phi(G, D) = arg \max_{s \in \Sigma} P(s|D, G), \quad (4)$$

where:

- D is the set of structured logs for a given session;
- G is the GenAI model parameterized through the system and user prompts;
- S is the resulting textual summary;
- Σ is the space of all possible textual outputs;
- $P(s \mid D, G)$ is the probability that a given sentence is the most faithful linguistic representation of the data.

The $arg\,max$ function returns the argument/sentence s for which the probability $P(s \mid D, G)$ is maximized. This formulation shows that the summary S is the result of optimizing the likelihood that the generated text expresses as fully and accurately as possible the structure of the data D given the context G.

During the experiments, the GenAI models used – ChatGPT and Claude – generated Markdown summaries with an average length of 4–6 sentences. The results showed that:

- 100% of the summaries were grammatically correct and easy to read;
- 91% correctly classified the behavior as normal or suspicious, consistent with the simulation scenario;
- 7% contained overly generalized conclusions (e.g., "activity appears normal" in cases of partial deviation);
- 2% were rejected by the system's validation due to missing the expected Markdown format.

These results confirm that the combination of stateless generation and automatic validation is an effective approach for achieving high accuracy without human intervention.

4.4. Behavioral Analysis Using Isolation Forest

For experimental demonstration, the Isolation Forest algorithm was used, trained on data generated by a GenAI script (with 15% anomalous records).

The results showed:

- Anomaly detection accuracy: approximately 94%;
- False positives (normal attempts classified as anomalies): below 6%;
- Strong dependence on the richness of behavioral data – longer sessions significantly increased accuracy.

The conclusion is that integrating the algorithm in real-time would allow the system to compute a probabilistic risk score for each session and trigger GenAI summaries with varying levels of detail.

5. Discussion

The developed model offers the following advantages:

- It transforms technical logs into concise and comprehensible GenAI summaries, supporting administrators and accelerating incident response.
- The reactive architecture (Redis Pub/Sub, TTL) ensures low latency and real-time processing.
- Integrating behavioral and geographical factors adds an additional layer of security.
- Experiments with the prototype confirm good performance and stability.

The main limitations include:

- Dependence on external GenAI services, which introduces risks related to privacy, latency, and resilience; transitioning to local models is a potential solution.
- The possibility that GenAI may generate variable or incorrect interpretations.
- Ethical and legal risks, including the need for GDPR compliance.
- The requirement for human verification in critical decisions.

Effective use of GenAI in security requires a clear separation of roles:

- GenAI analyzes and explains;
- The system reacts automatically within predefined boundaries;
- The human operator makes the final decision.

This three-level organization prevents excessive autonomy of artificial intelligence and ensures ethical and legal reliability in security processes.

In addition to improving the prototype, another direction in the development of Forest Gate is its integration with artificial neural networks. Based on 150,000 accumulated login events, experiments were conducted to predict likely geolocations for future access attempts. Using the Orange Data Mining environment, LSTM Autoencoder models were trained on normal login sequences for individual geo-clusters. The network learns to reconstruct typical login patterns, and deviations are interpreted as indicators of elevated risk.

After training on 150,000 samples, the model predicted the geolocations of the next 2,000 login attempts. The results show:

- Bulgaria: 226 expected attempts with approximately 77% probability of successful access (Figure 2).

- Seychelles: 35 attempts with approximately 6% probability of success.
- China/Hong Kong: 38 attempts with 0% predicted likelihood of successful access.

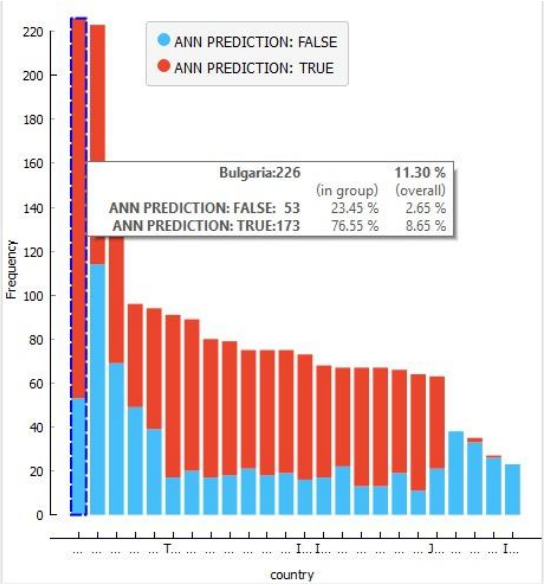


Figure 2. Probability of geolocations during login attempts according to the neural network.

The results can be visualized on a geographical map (Figure 3), which shows the global forecast of success or failure for the next 2,000 login attempts.

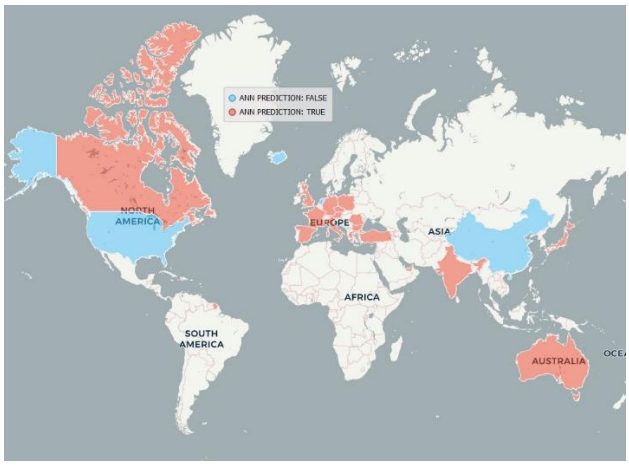


Figure 3. Geographical map of the predicted geolocations for the next 2,000 login attempts, according to the neural network.

6. Conclusion

This paper presented an experimental prototype of a secure authentication and user-session monitoring system built on the synergy of behavioral analysis, GenAI, and event-driven architecture. The Forest Gate project demonstrates that even at an early stage, it is possible to achieve effective, explainable, and adaptive access protection for web applications using lightweight and open technologies.

The experimental results show that GenAI can be integrated without degrading the performance or stability of authentication systems. The prototype demonstrates several advantages, including enhanced security, the transformation of technical data into comprehensible summaries, and accelerated detection and notification of critical events. At the same time, it must be acknowledged that the use of GenAI introduces risks related to the generation of inaccurate or hallucinated conclusions.

In this context, Forest Gate can be viewed as a next-generation hybrid platform that combines the power of language models with the predictability of engineered mechanisms. The approach is applicable not only in academic and experimental environments but also as a foundation for extensions into corporate or cloud-based solutions.

References

[1] Manyika, J., & Roxburgh, C. (2011). *The great transformer: The impact of the Internet on economic growth and prosperity*. McKinsey Global Institute, 1(0360-8581), 1-9.

[2] OECD. (2023). *Country digital education ecosystems and governance: A companion to Digital Education Outlook*. OECD Publishing. https://doi.org/10.1787/906134d4-en

[3] Liu, Y., & Qiang, Z. (2023). *Digital progress and trends report 2023* (English). World Bank Group. Retrieved from http://documents.worldbank.org/curated/en/099031924192524293

[4] Sahin, S., & Li, F. (2021). Don't forget the stuffing! Revisiting the security impact of typo-tolerant password authentication. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*. https://doi.org/10.1145/3460120.3484791

[5] Otta, S. P., et al. (2023). A systematic survey of multi-factor authentication for cloud infrastructure. *Future Internet*, *15*(4), 146. https://doi.org/10.3390/fi15040146

[6] Rabzelj, M., & Sedlar, U. (2025). Beyond the leak: Analyzing the real-world exploitation of stolen credentials using honeypots. *Sensors*, *25*(12), 3676. https://doi.org/10.3390/s25123676

[7] Matzen, A., et al. (2025). Challenges and potential improvements for passkey adoption—A literature review with a user-centric perspective. *Applied Sciences*, *15*(8), 4414. https://doi.org/10.3390/app15084414

[8] Verizon (DBIR). (2024). *Data breach investigations report*. Retrieved November 8, 2025, from https://www.verizon.com/business/resources/reports/2024-dbir-data-breach-investigations-report.pdf

[9] Baig, A. F., & Eskeland, S. (2021). Security, privacy, and usability in continuous authentication: A survey. *Sensors*, *21*(17), 5967. https://doi.org/10.3390/s21175967

[10] Verma, A., Moghaddam, V., & Anwar, A. (2022). Data-driven behavioural biometrics for continuous and adaptive user verification using smartphone and smartwatch. *Sustainability*, *14*(12), 7362. https://doi.org/10.3390/su14127362

[11] Papaioannou, M., et al. (2023). A survey on quantitative risk estimation approaches for secure and usable user authentication on smartphones. *Sensors*, *23*(6), 2979. https://doi.org/10.3390/s23062979

[12] Bonneau, J., et al. (2015). Passwords and the evolution of imperfect authentication. *Communications of the ACM*, *58*(7), 78–87. https://doi.org/10.1145/2699390

[13] Ometov, A., et al. (2018). Multi-factor authentication: A survey. *Cryptography*, *2*(1), 1. https://doi.org/10.3390/cryptography2010001

[14] Hasan, S. S., et al. (2025). A review on secure authentication mechanisms for mobile security. *Sensors*, *25*(3), 700. https://doi.org/10.3390/s25030700

[15] IBM. (2024). *Cost of a data breach report 2025: The AI oversight gap – Top insights, AI impact, and risk reduction best practices*. IBM. Retrieved November 8, 2025, from https://ibm.webcasts.com/starthere.jsp?ei=1726278

[16] Bułat, R., & Ogiela, M. R. (2023). Personalized context-aware authentication protocols in IoT. *Applied Sciences*, *13*(7), 4216. https://doi.org/10.3390/app13074216

[17] Ryu, R., Yeom, S., Herbert, D., & Dermoudy, J. (2023). A comprehensive survey of context-aware continuous implicit authentication in online learning environments. *IEEE Access*, *11*, 24561–24573. https://doi.org/10.1109/ACCESS.2023.3253484

[18] Saleem, J., et al. (2025). Machine learning-enhanced attribute-based authentication for secure IoT access control. *Sensors*, *25*(9), 2779. https://doi.org/10.3390/s25092779

[19] de Almeida, M. G., & Canedo, E. D. (2022). Authentication and authorization in microservices architecture: A systematic literature review. *Applied Sciences*, *12*(6), 3023. https://doi.org/10.3390/app12063023

[20] Sharif, A., et al. (2022). The eIDAS regulation: A survey of technological trends for European electronic identity schemes. *Applied Sciences*, *12*(24), 12679. https://doi.org/10.3390/app122412679

[21] Han, A. H., & Lee, D. H. (2023). Detecting risky authentication using the OpenID Connect token exchange time. *Sensors*, *23*(19), 8256. https://doi.org/10.3390/s23198256

[22] Venčkauskas, A., et al. (2023). Enhancing microservices security with token-based access control method. *Sensors*, *23*(6), 3363. https://doi.org/10.3390/s23063363

[23] Wang, Y., Zhang, X., & Hu, H. (2023). Continuous user authentication on multiple smart devices. *Information*, *14*(5), 274. https://doi.org/10.3390/info14050274

[24] Bansal, P., & Ouda, A. (2024). Continuous authentication in the digital age: An analysis of reinforcement learning and behavioral biometrics. *Computers*, *13*(4), 103. https://doi.org/10.3390/computers13040103

[25] Wang, S., et al. (2021). Machine learning in network anomaly detection: A survey. *IEEE Access*, *9*, 152379–152396. https://doi.org/10.1109/ACCESS.2021.3126834

[26] Kyaw, P. H., Gutierrez, J., & Ghobakhlou, A. (2024). A systematic review of deep learning techniques for phishing email detection. *Electronics*, *13*(19), 3823. https://doi.org/10.3390/electronics13193823

[27] Wilk-Jakubowski, J. L., et al. (2025). Machine learning and neural networks for phishing detection: A systematic review (2017–2024). *Electronics*, *14*(8), 3744. https://doi.org/10.3390/electronics14183744

[28] Balogh, Š. (2024). Using generative AI models to support cybersecurity analysts. *Electronics*, *13*(23), 4718. https://doi.org/10.3390/electronics13234718

[29] Lee, J., et al. (2025). LogRESP-Agent: A recursive AI framework for context-aware log anomaly detection and TTP analysis. *Applied Sciences*, *15*(13), 7237. https://doi.org/10.3390/app15137237

[30] Palma, G., et al. (2025). Leveraging large language models for scalable and explainable cybersecurity log analysis. *Journal of Cybersecurity and Privacy*, *5*(3), 55. https://doi.org/10.3390/jcp5030055

[31] Zhu, S., et al. (2022). An optimized algorithm for dangerous driving behavior identification based on unbalanced data. *Electronics*, *11*(10), 1557. https://doi.org/10.3390/electronics11101557

[32] Chen, J., et al. (2023). An anomaly detection method for wireless sensor networks based on the improved isolation forest. *Applied Sciences*, *13*(2), 702. https://doi.org/10.3390/app13020702

[33] Villegas-Ch., et al. (2023). IoT anomaly detection to strengthen cybersecurity in the critical infrastructure of smart cities. *Applied Sciences*, *13*(19), 10977. https://doi.org/10.3390/app131910977

[34] Chua, W., et al. (2024). Web traffic anomaly detection using isolation forest. *Informatics*, *11*(4), 83. https://doi.org/10.3390/informatics11040083

[35] Alboqmi, R., & Gamble, R. F. (2025). Enhancing microservice security through vulnerability-driven trust in the service mesh architecture. *Sensors*, *25*(3), 914. https://doi.org/10.3390/s25030914

[36] Mushtaq, S., Mohsin, M., & Mushtaq, M. M. (2025). A systematic literature review on the implementation and challenges of zero trust architecture across domains. *Sensors*, *25*(19), 6118. https://doi.org/10.3390/s25196118



https://csirjournal.com