

УДК 004.8:005.52

DOI: 10.31673/2412-9070.2025.027261

Я. Д. ЯЦЕНКО, аспірант;

ORCID: 0009-0001-2610-222X

О. В. ЖИДКА, старший викладач,

ORCID: 0009-0009-4272-9071

О. Я. КІС, аспірант,

ORCID: 0009-0006-9174-2588

Державний університет інформаційно-комунікаційних технологій, Київ

АНАЛІЗ СУЧАСНИХ СИСТЕМ ПРЕДИКТИВНОЇ АНАЛІТИКИ

У статті проведено порівняльний аналіз сучасних систем предиктивної аналітики (СПА), що використовують методи машинного навчання. Розглянуто ключові функціональні можливості, архітектурні особливості, рівень автоматизації та сферу застосування найпопулярніших платформ, зокрема IBM Watson Studio, Google Vertex AI, Microsoft Azure ML, Amazon SageMaker, RapidMiner, DataRobot, H2O.ai та SAS Predictive Analytics. Особливу увагу приділено класифікації СПА за типом задач, підходами до розгортання, а також інтеграційними можливостями з джерелами даних. Визначено сильні та слабкі сторони кожної системи, їхню продуктивність, зручність використання, відповідність стандартам безпеки та можливість масштабування. На основі проведеного аналізу запропоновано рекомендації щодо вибору СПА залежно від типу організації, рівня технічної підготовки користувачів та специфіки прикладних задач.

Ключові слова: предиктивна аналітика, машинне навчання, хмарні платформи, системи підтримки прийняття рішень, AutoML.

Вступ

Постановка проблеми. У зв'язку із стрімким зростанням обсягів даних та ускладненням бізнес-процесів, зростає потреба у високоточних інструментах, здатних не лише аналізувати історичну інформацію, а й прогнозувати майбутні події. Системи предиктивної аналітики, що базуються на методах машинного навчання, стали важливим компонентом цифрових трансформацій у різних галузях. Проте на сучасному ринку представлена велика кількість таких рішень, які відрізняються між собою за функціональністю, рівнем автоматизації, архітектурою, доступністю, вимогами до ресурсів та гнучкістю налаштування. Відсутність єдиного підходу до систематизації та порівняння цих платформ створює труднощі при виборі оптимального рішення для конкретних задач. У таких умовах виникає необхідність у проведенні структурованого порівняльного аналізу найпоширеніших сучасних СПА з урахуванням їхніх сильних і слабких сторін, щоб надати обґрунтовані рекомендації щодо їх вибору та впровадження.

Аналіз останніх досліджень. Упродовж останніх років у світовій науковій спільноті спостерігається зростаючий інтерес до систем предиктивної аналітики, що відображається у численних дослідженнях, присвячених їх впровадженню у різних галузях.

У вітчизняних джерелах значну увагу приділено теоретичним і прикладним аспектам впровадження предиктивної аналітики. Зокрема, Кашена Н. Б., Остапенко Р. М. та Велієва В. О. [1] проаналізували технологічні тренди в галузі штучного інтелекту, що створюють підґрунтя для формування ефективних СПА. Дмитришин Б. В. і Боровий М. В. [2] підкреслили зв'язок між впровадженням аналітичних інструментів і підвищенням конкурентоспроможності підприємств. Пшик-Ковальська О. О., Семенець М. Д. і Бруцяк Р. І. [3] зробили огляд програмних рішень з можливістю побудови прогнозів, включаючи як класичні методи статистичного аналізу, так і сучасні ML-підходи. Філіпова Л. Я. [4] акцентувала на важливості формування компетенцій аналітиків у сфері побудови прогнозних моделей.

Паралельно активно розвиваються дослідження за кордоном. Автори [5] зосередили увагу на можливостях застосування предиктивної аналітики для підвищення ефективності виявлення та реагування на кіберзагрози у режимі реального часу.

© Яценко Я. Д., Жидка О. В., Кіс О. Я., 2025

Автори [6] аналізують сучасні тенденції та майбутні напрями застосування великих даних у предиктивній аналітиці, охоплюючи сфери охорони здоров'я, фінансів, маркетингу та підкреслюють зростаюче впровадження технологій великих даних у цих секторах, що сприяє покращенню точності та ефективності предиктивних моделей.

Крім того, організація Gartner щорічно публікує аналітичні огляди ринку інструментів предиктивної аналітики (Magic Quadrant for Data Science and Machine Learning Platforms), у яких аналізуються лідери ринку (IBM, SAS, DataRobot, Microsoft, Amazon Web Services тощо), а також перспективи розвитку індустрії.

Мета роботи. Здійснити порівняльний аналіз сучасних систем предиктивної аналітики, охарактеризувати їх функціональні можливості, архітектурні підходи та сфери застосування, а також виявити перспективні напрями розвитку СПА в контексті інтеграції з технологіями машинного навчання та великих даних.

Основна частина

Класифікація сучасних систем предиктивної аналітики. Системи предиктивної аналітики – це програмно-апаратні комплекси або сервіси, що забезпечують побудову прогнозних моделей на основі історичних даних з використанням алгоритмів машинного навчання та статистики. Залежно від ключових характеристик, сучасні СПА можна класифікувати за кількома ознаками.

За типом задач:

- класифікація – передбачає розподіл об'єктів на категорії (наприклад, виявлення шахрайства, оцінка ризиків, сегментація клієнтів);
- регресія – використовується для передбачення числових значень, таких як попит, дохід або зміна вартості активів;
- прогнозування часових рядів – орієнтоване на моделювання динаміки процесів у часі (наприклад, обсяг продажів, навантаження на сервер, фінансові ринки).

Універсальні платформи часто підтримують усі три типи задач, але деякі СПА оптимізовані під конкретний тип застосування.

За ступенем автоматизації:

- AutoML-системи – повністю автоматизовані рішення, які автоматично обирають моделі, гіперпараметри та виконують підготовку даних (наприклад, Google Cloud AutoML, H2O Driverless AI);
- напівавтоматизовані рішення – користувач може втручатися на окремих етапах побудови моделі, комбінуючи власні знання з автоматизованими підказками (наприклад, Azure Machine Learning, DataRobot);
- кастомні рішення – платформи або фреймворки, що потребують повного ручного управління та програмування (наприклад, Scikit-learn, TensorFlow, PyTorch) [7].

Ступінь автоматизації визначає рівень експертизи, необхідної для ефективного використання СПА.

За підходами до розгортання:

- он-преміс-рішення (On-premises) – системи, що розгортаються на локальних серверах організації; переважно використовуються у середовищах з підвищеними вимогами до безпеки та контролю;
- хмарні рішення (Cloud-based) – пропонують гнучкість, масштабованість та доступність (наприклад, Amazon SageMaker, Google Vertex AI, Microsoft Azure ML);
- гібридні моделі – поєднують локальні обчислення з хмарною аналітикою, дозволяючи досягти балансу між контролем над даними та обчислювальними потужностями.

Таким чином, класифікація СПА за цими трьома ознаками дозволяє обрати оптимальне рішення відповідно до бізнес-задач, технічних вимог та ресурсів організації.

Критерії порівняння систем предиктивної аналітики. Для об'єктивного оцінювання ефективності сучасних СПА доцільно враховувати низку технічних, функціональних та корис-

тувацьких критеріїв. Нижче подано ключові параметри порівняння, що допомагають визначити доцільність використання тієї чи іншої системи у певному контексті.

1. Архітектура. Важливим фактором є архітектура платформи: наскільки вона модульна (можливість заміни чи оновлення окремих компонентів), масштабована (підтримка зростання обсягів даних і користувачів), а також наявність API (інтерфейсів для інтеграції з іншими системами). Гнучка архітектура дозволяє адаптувати СПА до змін бізнес-середовища.

2. Підтримувані алгоритми ML/AI. СПА варто оцінювати за широтою підтримки алгоритмів машинного навчання та штучного інтелекту від класичних моделей (лінійна регресія, дерева рішень) до сучасних методів (нейронні мережі, ансамблеві моделі, глибоке навчання). Платформи з вбудованими механізмами AutoML значно полегшують процес вибору і налаштування моделей [8].

3. Інтеграція з джерелами даних. Ефективність СПА значною мірою залежить від інтеграційних можливостей: підтримка підключення до баз даних (SQL, NoSQL), сховищ у хмарі (AWS S3, Google Cloud Storage), потокових сервісів (Kafka, Azure Event Hub) та API сторонніх систем (ERP, CRM тощо).

4. Інструменти підготовки та візуалізації даних. Наявність інструментів для очищення, трансформації та нормалізації даних, а також вбудованих засобів візуалізації (графіки, дашборди, heatmap тощо) забезпечує зручність попередньої обробки даних і покращує інтерпретацію результатів моделювання.

5. Продуктивність і точність. Продуктивність системи оцінюється за часом обробки запитів і побудови моделей, а точність – за метриками моделювання (RMSE, ROC-AUC, precision/recall тощо). Сучасні СПА мають механізми оптимізації продуктивності за рахунок паралельних обчислень і GPU-прискорення.

6. Безпека та відповідність стандартам. Критичним є рівень безпеки даних, включаючи шифрування, керування доступом, аудит, а також відповідність міжнародним стандартам (GDPR, ISO/IEC 27001). Це особливо актуально для рішень у медичній, фінансовій та державній сферах [9].

7. Юзер-френдлі інтерфейс, документація, підтримка. Зручність використання системи визначається якістю інтерфейсу користувача, доступністю документації, навчальних матеріалів та наявністю технічної підтримки (форумів, чат-ботів, служби підтримки). СПА, орієнтовані на бізнес-користувачів, повинні мати інтуїтивно зрозумілий інтерфейс та мінімальні вимоги до програмної підготовки користувача.

Ці критерії дозволяють провести ґрунтовний аналіз і сформувані об'єктивну оцінку систем у процесі вибору СПА для конкретного середовища чи проєкту.

Огляд сучасних платформ предиктивної аналітики. IBM Watson Studio – потужна платформа для розробки моделей машинного навчання та штучного інтелекту. Підтримує Jupyter Notebooks, RStudio, AutoAI. Має широкий набір інструментів для обробки даних, візуалізації, MLOps. Забезпечує інтеграцію з IBM Cloud, підтримує хмарне і локальне розгортання. Орієнтована на підприємства, що працюють з критично важливою інформацією.

SAS Predictive Analytic – один з лідерів у сфері аналітики з багаторічним досвідом. Його платформа підтримує розширене статистичне моделювання, візуальну аналітику та прогнозування часових рядів. Має надійний захист даних, сертифікований за стандартами безпеки. Орієнтована на великі організації з потребою у точних, стандартизованих аналітичних звітах.

Google Cloud Vertex AI – хмарна платформа Google для повного циклу ML, що поєднує AutoML і кастомне моделювання. Має інтеграцію з BigQuery, Dataflow, Looker Studio. Підтримує MLOps, CI/CD для моделей. Платформа відзначається високою продуктивністю, гнучкістю і зручністю розгортання [10].

Microsoft Azure Machine Learning Studio – гнучка хмарна платформа з підтримкою AutoML, візуального програмування, Python SDK. Платформа дозволяє створювати, тренувати, тестувати та розгортати ML-моделі в інтегрованому середовищі. Має велику кількість шаблонів, зручну інтеграцію з Power BI та іншими сервісами Microsoft [11].

Amazon SageMaker – одна з наймасштабованіших хмарних платформ для ML. Пропонує інструменти для підготовки даних, AutoML, тренування моделей на GPU, розгортання в режимі реального часу. Сильна сторона – продуктивність, гнучке налаштування інфраструктури, підтримка великих обсягів даних [12].

RapidMiner – інструмент з інтуїтивним інтерфейсом drag-and-drop, орієнтований на користувачів без глибоких знань у програмуванні. Підтримує понад 1500 алгоритмів та операцій, зручну інтеграцію з базами даних та BI-інструментами. Платформа має хмарну та локальну версії, популярна в освіті та серед SMB [13].

DataRobot – спеціалізована платформа AutoML, яка забезпечує автоматизоване тренування моделей, пояснення рішень (explainable AI), інструменти MLOps. Має високий рівень автоматизації, вбудовані рекомендації для бізнес-користувачів. Використовується в галузях, де потрібна швидка аналітика (фінанси, охорона здоров'я) [14].

H2O.ai – високопродуктивна платформа з відкритим вихідним кодом для розробників і дослідників. Підтримує AutoML, глибоке навчання, GPU-прискорення, а також Spark. Відома своєю швидкістю обробки великих обсягів даних, активною спільнотою та прозорими алгоритмами.

Порівняння систем предиктивної аналітики за наведеними вище критеріями

Критерій	IBM Watson Studio	SAS Predictive Analytics	Google Vertex AI	Azure ML Studio	Amazon SageMaker	RapidMiner	DataRobot	H2O.ai
Архітектура	Модульна, API, хмара/локал	Моноліт, API, локал/хмара	Хмара, API, Kubeflow	Хмара, API, інтегр. з Power BI	Хмара, API, модульна	Локал/хмара, drag-drop	Хмара, AutoML-платформа	Локал/хмара, open-source
AutoML/Алгоритми ML	AutoAI, sklearn, XGBoost	Багатий набір, без AutoML	AutoML + TensorFlow PyTorch	AutoML + кастомні моделі	AutoML, кастомні моделі	>1500 модулів, AutoML	AutoML + explainable AI	AutoML, GBM, DL, XGBoost
Інтеграція з джерелами даних	DB2, Hadoop, AWS, FTP	Oracle, Hadoop, DB2	BigQuery, GCS, API	SQL, Azure DB, API	S3, Redshift, Snowflake	SQL, CSV, NoSQL	API, SQL, BI-системи	HDFS, S3, SQL
Підготовка/візуалізація даних	SPSS, Cognos, графіки	SAS Visual Analytics	Looker, Data Studio	Power BI, інтегровано	SageMaker Studio	Візуальний pipeline	Вбудована аналітика	Інтеграція з Tableau/R
Продуктивність/точність	Висока, підтримка GPU	Стабільна класична точність	Висока, автоматичне масштабування	Середня-висока, GPU	Висока, розподілені обчислення	Середня, залежить від сценарію	Висока, оптимізація моделей	Висока, швидка обробка
Безпека/стандарти	ISO, GDPR, HIPAA	Повна сертифікація	GDPR, ISO/IEC 27001	ISO, SOC, FedRAMP	GDPR, HIPAA, SOC	Базова, залежить від версії	GDPR, Explainability focus	Open-source + enterprise
Інтерфейс/документація/підтримка	UI, багата док., підтримка	Професійна підтримка	Зручний UI, розширена док.	Інтуїтивний, підтримка Azure	Інтеграція з AWS Console	Drag&drop, підтримка форумів	Бізнес-орієнтований UI	Спільнота, техпідтримка

Узагальнення результатів порівняння СПА. Розглянувши зазначені вище критерії, можна зробити наступні узагальнення. Для корпоративного сектору (Enterprise) найкраще підходять IBM Watson Studio (багатофункціональність, безпечне розгортання), SAS Predictive Analytics (високий рівень довіри, регламентованість процесів), Azure ML Studio (гнучка інтеграція з BI-середовищем Microsoft). Для масштабованих cloud-рішень варто використовувати Google Vertex AI та Amazon SageMaker (найбільша гнучкість, поєднання AutoML і кастомізації, а також потужна підтримка DevOps-практик). Для швидкого старту без програмування рекомендовано RapidMiner (drag&drop інтерфейс) та DataRobot (бізнес-орієнтована AutoML платформа, не вимагає знань у сфері машинного навчання). Для дослідницьких/наукових цілей або

стартапів підходить H2O.ai (безкоштовна система з відкритим кодом, високою продуктивністю та активною спільнотою користувачів).

Висновки

Проведений аналіз сучасних СПА засвідчив їхню різноманітність за функціональністю, архітектурою, рівнем автоматизації та зручністю використання. Залежно від потреб, користувачі можуть обирати між потужними корпоративними платформами, масштабованими хмарними рішеннями, інструментами для швидкого старту без програмування або відкритими системами для наукових досліджень і стартапів.

У перспективі СПА розвиватимуться у напрямку подальшої автоматизації, розширення функцій AutoML, інтеграції з no-code платформами, підвищення прозорості моделей та забезпечення етичності аналізу. Водночас актуальним залишається питання доступності таких рішень для малого та середнього бізнесу. Подальші дослідження доцільно зосередити на оцінюванні ефективності СПА в умовах змінних даних, вивченні етичних викликів у чутливих сферах та створенні гібридних аналітичних моделей, що поєднують предиктивні, дескриптивні й прескриптивні підходи.

Список літератури

1. Kashchena N., Ostapenko R., Veliieva V. *Business intelligence as a data processing tool // Economy and Society*. – 2024. – No. 62. – <https://doi.org/10.32782/2524-0072/2024-62-14>
2. Dmytryshyn B. V., Borovyi M. V. *Business analytics and its role in managing the competitiveness of an enterprise // Central Ukrainian Scientific Bulletin. Economic Sciences*. – 2020. – No. 5(38). – P. 214–220. – http://nbuv.gov.ua/UJRN/Npkntu_e_2020_5_24
3. Pshyk-Kovalska O. O., Semenets M. D., Brutsiak R. I. *Business intelligence tools to improve business operations // Efficient Economy*. – 2024. – No. 8. – <https://doi.org/10.32702/2307-2105.2024.8.81>
4. Filipova L. Ya. *Business intelligence systems: Current trends in development // Library Science. Document Science. Informology*. – 2022. – No. 1. – P. 43–48. – <https://doi.org/10.32461/2409-9805.1.2022.257258>
5. Rassam M.A., Zainal A., Maarof M.A., Al-Qershi O.M., Mohd Sahran S. *Enhancing cyber security through predictive analytics: Real-time threat detection and response // arXiv*. – 2024. – [arXiv:2407.10864](https://doi.org/10.48550/arXiv.2407.10864). – <https://doi.org/10.48550/arXiv.2407.10864>
6. Jasim W.A., Alnajar H.R., Hamid A.S., Aldabagh D.A., Shabala Y. *The Role of Big Data in Predictive Analytics: Current Trends and Future Directions // Journal of Ecohumanism*. – 2024. – Vol. 3, No. 5. – P. 422–443. – DOI: <https://doi.org/10.62754/joe.v3i5.3915>
7. Tableau. *A Guide to Predictive Analytics // Tableau Software*. – 2024. – <https://www.tableau.com/analytics/what-is-predictive-analytics>
8. Boiko V. *Predictive analytics tools: a complete guide with examples // Itransition*. – 2025. – <https://www.itransition.com/predictive-analytics/tools>
9. *Deep Dive into Predictive Analytics Models and Algorithms // Maruti Techlabs*. – 2024. – <https://marutitech.com/predictive-analytics-models-algorithms/>
10. Google Vertex AI. – [Електронний ресурс]. – Режим доступу: <https://cloud.google.com/vertex-ai>
11. Microsoft Azure ML. – [Електронний ресурс]. – Режим доступу: <https://azure.microsoft.com/services/machine-learning>
12. Amazon SageMaker. – [Електронний ресурс]. – Режим доступу: <https://aws.amazon.com/sagemaker>
13. RapidMiner. – [Електронний ресурс]. – Режим доступу: <https://rapidminer.com>
14. DataRobot. – [Електронний ресурс]. – Режим доступу: <https://www.datarobot.com>

D. Yatsenko, O. Zhydka, O. Kis

ANALYSIS OF MODERN PREDICTIVE ANALYTICS SYSTEMS

Abstract: The article presents a comparative analysis of modern predictive analytics systems (PAS) that employ machine learning methods. The study focuses on evaluating the core capabilities, architectural models, and application areas of widely used platforms, including IBM Watson Studio, Google Vertex AI, Microsoft Azure Machine Learning Studio, Amazon SageMaker, RapidMiner, DataRobot, H2O.ai, and SAS Predictive Analytics. Special attention is given to the classification of PAS by the types of tasks they solve (classification, regression, time series forecasting), levels of automation (AutoML, semi-automated, and custom solutions), and deployment models (on-premise, cloud-based, and hybrid).

The analysis highlights key criteria for comparison: architectural flexibility and scalability, support for machine learning algorithms and AutoML features, integration with diverse data sources, data preparation and visualization tools, model performance and accuracy, security compliance (GDPR, ISO), and user experience including interface convenience, documentation, and vendor support.

Each system is assessed in terms of its strengths and weaknesses, depending on its suitability for various use cases. Enterprise-level platforms such as IBM Watson Studio, SAS, and Azure ML Studio are identified as optimal for regulated environments and large-scale deployments. In contrast, cloud-native solutions like Google Vertex AI and Amazon SageMaker demonstrate high flexibility and DevOps integration, making them ideal for scalable AI projects. Tools such as RapidMiner and DataRobot are more suitable for rapid prototyping and business users due to their intuitive interfaces and AutoML capabilities. Open-source platforms like H2O.ai are shown to be effective for research, experimentation, and startups thanks to their performance, transparency, and community support.

The paper concludes by outlining the current trends in PAS development and identifying promising directions for future research. These include the further integration of explainable AI (XAI), ethical data handling, and the convergence of predictive analytics with real-time data processing in cloud ecosystems. Based on the analysis, practical recommendations are proposed to assist organizations in selecting appropriate PAS tools aligned with their technical needs and strategic objectives.

Keywords: predictive analytics, machine learning, cloud platforms, decision support systems, AutoML.
