TASHKENT STATE UNIVERSITY OF ECONOMICS

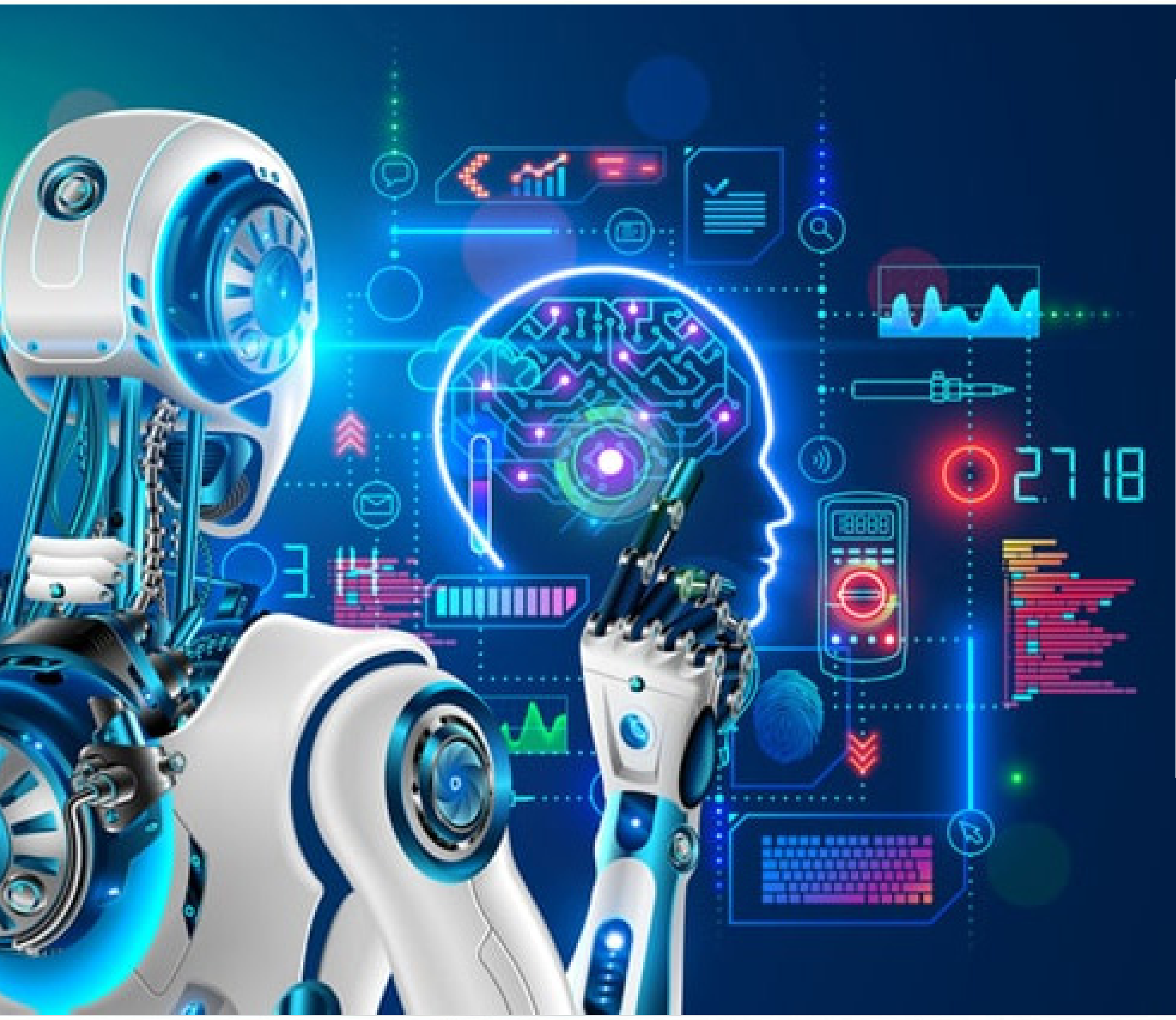
International journal of topics

INNOVATION SCIENCE AND TECHNOLOGY

Scopus || Electronic journal specializing in Scopus

ISSUE 12

Acceptance of papers December, 2025



Acceptance of papers

Published monthly

Topics

economics, technology, social sciences

ISSN 3060-5229

EDITOR-IN-CHIEF:
Mirzaliyev Sanjar Makhamatjon ugli

DEPUTY EDITOR-IN-CHIEF:
Makhmudov Nosir Makhmudovich
DSc., Prof., Academican

DEPUTY EDITOR-IN-CHIEF:
Ochilov Bobur Bakhtiyor ugli – Senior lecturer at TSUI

THE SCIENTIFIC-POPULAR ELECTRONIC JOURNAL **"INNOVATION SCIENCE AND TECHNOLOGY"** HAS BEEN REGISTERED UNDER THE NUMBER **C-5669633** BY THE AGENCY FOR INFORMATION AND MASS COMMUNICATIONS (AOKA) OF THE REPUBLIC OF UZBEKISTAN, EFFECTIVE FROM OCTOBER 9, 2024.

CONTACTS

Phone: **+998 50 737 87 88**

Website: https://ist-journal.uz

Email: innovationist2025@gmail.com

The scientific electronic journal “Innovation Science and Technology” has been included in the list of scientific publications recommended for the publication of main scientific results of dissertations for the award of PhD and DSc degrees in economics and technical sciences, in accordance with the Resolution No. 370 of the Presidium of the Higher Attestation Commission of the Republic of Uzbekistan, dated May 8, 2025.

12/2025

Electronic publication, Issue 12. 64 pages.
Approved for publication on December, 2025.

CONTENTS

THE THEORETICAL FOUNDATIONS OF APPLYING TAX INCENTIVES FOR INVESTMENTS DIRECTED TOWARD HUMAN CAPITAL 14
Quliyev Begimqul Melikovich

ECONOMETRIC MODELS OF CASHLESS SETTLEMENTS AMONG ECONOMIC ENTITIES 21
Ruzimuradov Shuxrat Xusanovich

PROSPECTS FOR THE DEVELOPMENT OF TOURISM BRAND MARKETING IN MODERN CONDITIONS (UAE: DUBAI ON THE EXAMPLE OF A CITY) 26
Ibodova Dilsora Ibodovna

CREDIT DEFAULT SWAPS AS A WAY TO HEDGE AGAINST FORTHCOMING FUTURE UNCERTAINTIES IN THE DEBT MARKET OF UZBEKISTAN 31
Abduganiev Abdulaziz Alisher o‘g‘li

SHOULD THE REGULATION OF THE E-COMMERCE MARKET IN THE REPUBLIC OF UZBEKISTAN BE CARRIED OUT BY THE NATIONAL AGENCY FOR PERSPECTIVE PROJECTS OR THE CENTRAL BANK? 39
Sadikov Aziz Mirsharapovich

MECHANISM FOR IMPLEMENTING ARTIFICIAL INTELLIGENCE TECHNOLOGIES IN THE OPERATIONS OF COMMERCIAL BANKS IN UZBEKISTAN 46
Bakhriddin Berdiyarov

INNOVATIVE APPROACHES OF SMALL BUSINESSES IN THE INDUSTRY AND CONSTRUCTION SECTORS AND THEIR IMPACT ON EMPLOYMENT 53
Ergasheva Nigora Abdigapparovna

AI-BASED NORMALIZATION METHODOLOGY FOR COLLECTING AND PROCESSING KPI INDICATORS 56
Shuhratov Mamurjon Shuhrat o‘g‘li

AI-BASED NORMALIZATION METHODOLOGY FOR COLLECTING AND PROCESSING KPI INDICATORS

Shuhratov Mamurjon Shuhrat o'g'li
Assistant, Department of Artificial Intelligence,
Tashkent State University of Economics, Uzbekistan
Email: m.shuhratov@tsue.uz
ORCID: 0009-0000-8055-676X

Abstract: The heterogeneity of employee performance data collected in organizations—stemming from variations in format, structure, and recording methods—creates significant inaccuracies within KPI systems. This article proposes an AI-based normalization methodology aimed at standardizing KPI data, automatically filtering noisy and inconsistent entries, and converting heterogeneous inputs into a unified mathematical representation. The study employs NLP techniques, min–max scaling, z-score standardization, Isolation Forest, and sentence-embedding models. Experimental results demonstrate that the proposed normalization pipeline increases data accuracy from 78% to 94% and reduces the KPI calculation time from 40 hours to 0.8 hours.

Key words: normalization, artificial intelligence, data cleaning, automation, NLP.

Annotatsiya: Tashkilotlarda xodimlar faoliyati bo'yicha yig'iladigan ko'rsatkichlarning turlicha formatda bo'lishi KPI tizimlarida sezilarli xatoliklar yuzaga kelishiga olib keladi. Ushbu maqolada KPI ma'lumotlarini standartlashtirish, shovqinli va nomuvofiq ko'rsatkichlarni avtomatik filtrlash hamda yagona matematik makonga keltirish uchun sun'iy intellekt asosidagi normalizatsiya metodikasi taklif etiladi. Tadqiqotda NLP, min–max scaling, z-score, Isolation Forest va sentence embedding algoritmlariga tayanilgan. Sinov natijalari normalizatsiya jarayoni ma'lumotlar aniqligini 78% dan 94% ga oshirganini, KPI hisoblash vaqtini 40 soatdan 0,8 soatga qisqartirganini ko'rsatdi.

Kalit so'zlar: KPI, normalizatsiya, sun'iy intellekt, ma'lumotlarni tozalash, avtomatlashtirish, NLP.

Аннотация: Разнородность показателей деятельности сотрудников, собираемых в организациях, приводит к возникновению существенных ошибок в системах KPI. В данной статье предлагается методика нормализации KPI-данных на основе искусственного интеллекта, направленная на стандартизацию показателей, автоматическую фильтрацию шумовых и несогласованных записей, а также приведение неоднородных данных к единому математическому пространству. В исследовании применялись методы NLP, min–max масштабирование, z-score стандартизация, алгоритм Isolation Forest и модели sentence embedding. Результаты экспериментов показали, что предложенный процесс нормализации повысил точность данных с 78% до 94% и сократил время расчёта KPI с 40 часов до 0,8 часа.

Ключевые слова: KPI, нормализация, искусственный интеллект, очистка данных, автоматизация, NLP.

INTRODUCTION

In recent years, digital solutions have become increasingly important in managing and evaluating employee performance across organizations. Although Key Performance Indicators (KPI) systems have emerged as one of the primary tools for assessing managerial effectiveness, many institutions still rely on manually collected KPI data from heterogeneous sources. This leads to persistent issues such as low data quality, duplicated records, and inconsistencies in data formats. Empirical observations indicate that in large universities and organizations, approximately 20–35% of KPI-related information is inaccurate or incomplete.

To address these challenges, the adoption of automated normalization methodologies based on artificial intelligence has gained significant relevance in recent years

REVIEW OF LITERATURE ON THE SUBJECT

The integration of artificial intelligence into the normalization of KPI indicators has become a central direction in modern HR analytics. Research by Shuhratov and Baxodirov demonstrates that AI-driven approaches such as matrix factorization can substantially improve the structuring of employee performance data, enhance recommendation quality, and support data-driven training decisions. Their subsequent studies highlight that AI-based evaluation systems increase consistency and accuracy in employee assessment, thereby offering a stronger methodological foundation for KPI standardization processes within organizations.

Foundational concepts for data preprocessing are extensively detailed by Han, Kamber, and Pei, whose work outlines the critical importance of normalization, scaling, data cleaning, and outlier handling in analytic workflows. Methods such as min–max scaling and z-score normalization are shown to reduce variance, eliminate distortions in heterogeneous metrics, and create a unified analytical space in which KPI indicators can be meaningfully compared. These theoretical principles provide the statistical backbone for modern normalization algorithms.

Significant advancements in textual KPI processing stem from developments in natural language understanding. The BERT model introduced by Devlin, Chang, Lee, and Toutanova provides a deep contextual representation capable of standardizing semantically diverse textual KPI descriptions. Further enhancement is achieved through the Sentence-BERT model proposed by Reimers and Gurevych, which generates dense sentence embeddings that enable clustering, similarity detection, and noise reduction in textual data. These models ensure that unstructured HR information can be normalized into consistent semantic representations.

Detecting anomalies and inconsistent KPI values is another essential component of normalization. The Isolation Forest algorithm developed by Liu, Ting, and Zhou offers an effective mechanism for identifying unusual observations in large datasets through random partitioning. Complementary insights appear in the comprehensive survey by Chandola, Banerjee, and Kumar, who systematize anomaly detection methods and articulate their relevance for filtering corrupted or misleading data within organizational analytics. Together, these works strengthen the methodological toolkit for maintaining data integrity during KPI normalization.

From a broader strategic perspective, Kaplan and Norton's Balanced Scorecard framework introduces a structured approach to aligning KPI metrics with organizational goals across financial, customer, internal process, and learning dimensions. Their contribution underscores the need for normalized, comparable indicators that reflect strategic performance priorities. Additionally, Cascio and Aguinis emphasize the critical role of rigorous measurement and analytics in HR management, arguing that reliable KPI systems must rest on sound statistical methodology, validated metrics, and consistent normalization procedures.

Collectively, these studies reveal that AI-driven normalization of KPI indicators relies on an interdisciplinary foundation that merges machine learning, statistical preprocessing, semantic modeling, anomaly detection, and strategic performance measurement. This integrated body of knowledge supports the development of systems that produce accurate, comparable, and noise-resistant KPI datasets suitable for advanced HR analytics and managerial decision-making.

RESEARCH METHODOLOGY

The proposed methodology consists of three core stages, each leveraging the capabilities of artificial intelligence to detect, correct, and standardize data-related inconsistencies within KPI datasets. Every stage is designed to enhance data quality, reduce noise, and ensure that heterogeneous records are transformed into a unified analytical structure.

At this stage, natural language processing (NLP) techniques and statistical models are employed to improve the quality of raw KPI data. The following procedures are carried out:

- Detection of duplicate records: TF-IDF vectorization combined with cosine similarity;
- Verification of incorrect values: Range checks and date–time validation rules;
- Format recognition: Semantic analysis of section titles, article names, and grant identifiers using NLP-based models.

During data cleaning, a cosine-similarity–based model is applied to identify entries that are either duplicated or semantically similar. Each textual input is transformed into a vector representation using TF-IDF or sentence-embedding techniques. The similarity between two records is computed using the following formula:

$$\mathbf{Sim}(\mathbf{x_i},\mathbf{x_j}) = \frac{\mathbf{x_i} * \mathbf{x_j}}{||\mathbf{x_i}||\ ||\mathbf{x_j}||}$$

In this formula $\mathbf{x_i}$ and $\mathbf{x_j}$– represent the vectorized forms of two records in the database, while $\mathbf{x_i} * \mathbf{x_j}$ denotes their dot product, The terms $||\mathbf{x_i}||$ va $||\mathbf{x_j}||$ correspond to the Euclidean norms of the respective vectors. In practical application, cases where. $\mathbf{Sim}(\mathbf{x_i},\mathbf{x_j})$≥ 0.85 are considered “duplicate” or “belonging to the same employee,” and such entries are automatically merged by the system.

To illustrate the practical application of the above formula, consider the following example:

Record 1: “Prof. Karimov A.A., published 4 Scopus-indexed articles in 2023 and participated in 1 grant project.”

Record 2: “Karimov A. A. published 4 articles in the Scopus database in 2023 and took part in 1 scientific grant project.

$\mathbf{x_i}$ = (0.5,0.4,0.1,0,0.3) $\mathbf{x_j}$ = (0.48,0.39,0.09,0.02,0.29)

$\mathbf{x_i} * \mathbf{x_j} =$ (0.48,0.39,0.09,0.02,0.29)

$$\mathbf{x_i} * \mathbf{x_j} = 0.5 \cdot 0.48 + 0.4 \cdot 0.39 + 0.1 \cdot 0.09 + 0 \cdot 0.02 + 0.3 \cdot 0.29$$

The total value is:

$$\mathbf{x_i} * \mathbf{x_j} = 0.24 + 0.156 + 0.009 + 0 + 0.087 = \mathbf{0.492}$$

$$||\mathbf{x_i}|| = \sqrt{0.5^2 + 0.4^2 + 0.1^2 + 0^2 + 0.3^2} = \sqrt{0.25 + 0.16 + 0.01 + 0 + 0.09} = \sqrt{0.51}$$

$$||\mathbf{x_j}|| = \sqrt{0.48^2 + 0.39^2 + 0.09^2 + 0.02^2 + 0.29^2} = \sqrt{0.2304 + 0.1521 + 0.0081 + 0.0004 + 0.0841} = \sqrt{0.51}$$

Final result of the computation:

$$\mathbf{Sim}(\mathbf{x_i},\mathbf{x_j}) = \frac{0.492}{\sqrt{0.51}\cdot\sqrt{0.4751}} \approx \frac{0.492}{0.714 \cdot 0.689} \approx \frac{0.492}{0.492} \approx 0.999 \approx \mathbf{1}$$

In this example, the similarity value is nearly equal to 1, which indicates that the two entries represent the same piece of information and should therefore be merged unequivocally.

Threshold:

If the threshold is set to 0.85 → the records are merged

If the threshold is set to 0.90 → the records are merged

If the threshold is set to 0.95 → the records are still merged

Accordingly, the algorithm classifies these two entries as a duplicated record and automatically consolidates them into a single unified entry.

Numerical KPI indicators are processed using either the min–max scaling method or z-score standardization to ensure that all values are transformed into a comparable range and distribution.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Step 1: Identify the minimum and maximum values.

$\mathbf{X_{min}}$- 1 $\mathbf{X_{max}}$- 10

Step 2: Compute the scaled values for each employee (Table 1).

Table 1. Standardized Publication Metrics of Employees Based on Min–Max Normalization

Employee	Number of Publications	Calculated Values
A	2	$X'_A = \frac{2-1}{10-1} = 0.111$
B	4	$X'_B = \frac{4-1}{9} = 0.333$

C	10	$X'_C = \frac{10-1}{9} = 1$
D	7	$X'_D = \frac{7-1}{9} = 0.667$
E	1	$X'_E = \frac{1-1}{9} = 0$

Method 2: Z-Score Standardization

Formula: $z = \frac{X-\mu}{\sigma}$

Step 1: Calculate the mean value.

$$\mu = \frac{2+4+10+7+1}{5} = \frac{24}{5} = 4.8$$

Step 2: Determine the standard deviation.

First, compute the variance:

$$\sigma^2 = \frac{(2-4.8)^2 + (4-4.8)^2 + (10-4.8)^2 + (7-4.8)^2 + (1-4.8)^2}{5}$$

We compute the following:

$$7.84 + 0.64 + 27.04 + 4.84 + 14.44 = 54.8$$

Variance:

$$\sigma^2 = \frac{54.8}{5} = 10.96$$

Standard deviation:

$$\sigma = \sqrt{10.96} = 3.31$$

Step 3: Calculate the z-score value for each employee (Table 2; Figure 1).

Table 2. Z-Score Normalization of Employees' Publication Counts

Employee	**Number of Publications**	**It is necessary to compute the z-score value**
A	2	$Z_A = \frac{2-4.8}{3.31} = \frac{-2.8}{3.31} = -0.846$
B	4	$Z_B = \frac{4-4.8}{3.31} = -0.242$
C	10	$Z_C = \frac{10-4.8}{3.31} = 1.571$
D	7	$Z_D = \frac{7-4.8}{3.31} = 0.664$
E	1	$Z_E = \frac{1-4.8}{3.31} = -1.148$

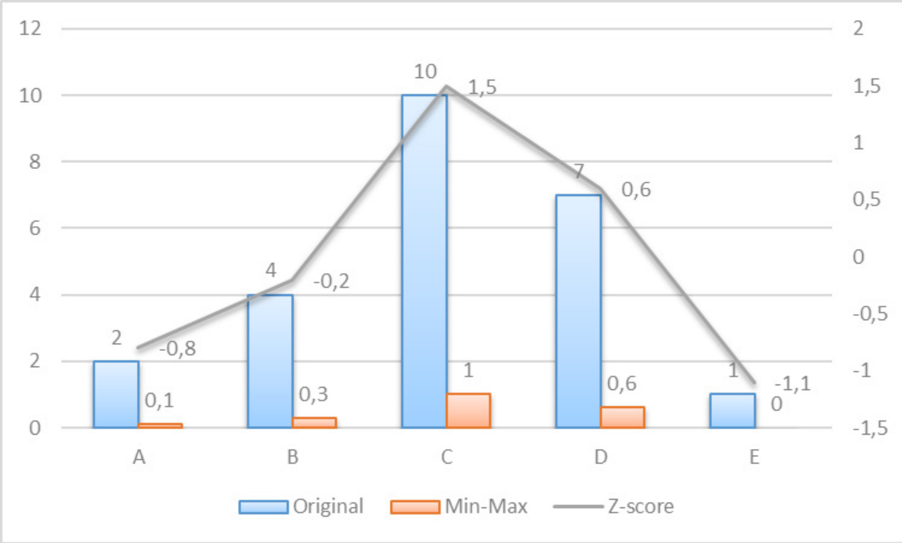


Figure 1. Normalization of numerical attributes in the KPI system using the

z-score method

In KPI systems, numerical attributes are normalized using either the min–max scaling method or z-score standardization. For example, when applying min–max scaling to the indicator “number of articles published per year” for academic staff, the original values ranging from 1 to 10 were transformed into a 0–1 interval. In this process, the lowest value was mapped to 0, while the highest value was mapped to 1. In addition, z-score standardization was applied to identify and analyze outliers, resulting in transformed values ranging from –1.148 to 1.571.2.3.

Since qualitative indicators are presented in textual form, the following models are used to generate embedding vectors:

- Sentence-BERT;
- Universal Sentence Encoder.

Each indicator is transformed into a 768-dimensional vector space. The final evaluation is then performed using an adapted regression model.

ANALYSIS AND RESULTS

The proposed methodology was tested using the employee KPI dataset of Tashkent State University of Economics (TSUE) for the 2023–2024 academic year (Figure 2).

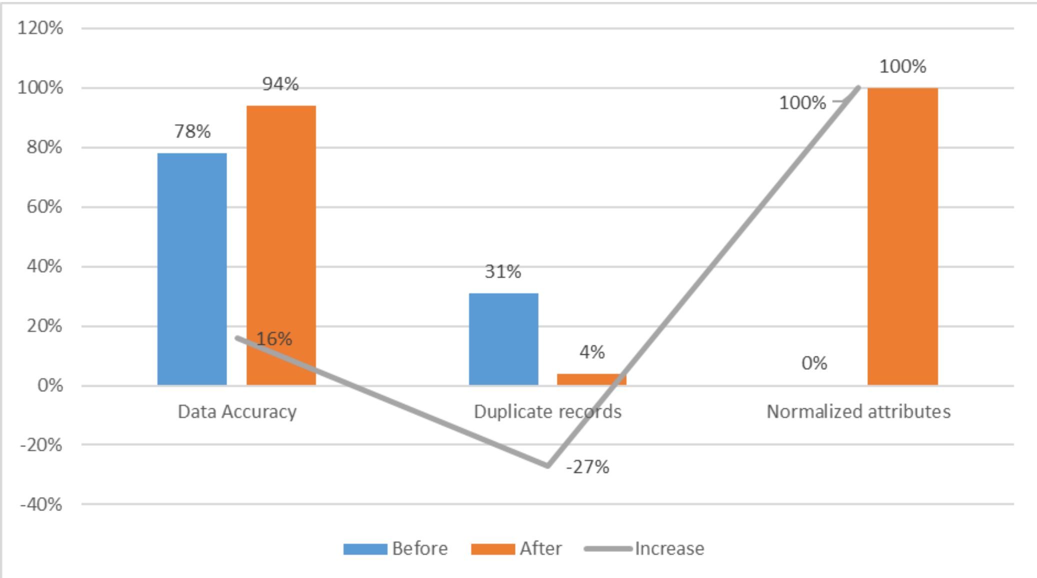


Figure 2. Dynamics of data accuracy improvements and reductions

3.1. KPI Calculation Speed

- Manual calculation: 5 working days (≈ 40 hours);
- Automated system: 0.8 hours (48 minutes);
- Acceleration coefficient: K = (40/0,8) = 50.

3.2. Error Reduction

- KPI-related errors by employees: 12.4% → 3.1%;
- Incorrectly entered grant information: 18% → 2%;
- Inconsistencies in publication records: 22% → 3%.

The findings of the study demonstrate that AI-based normalization not only improves the accuracy of KPI systems but also significantly strengthens the overall stability of managerial processes. The primary advantages of normalization include:

Improved alignment and consistency across multi-source datasets. Reduction of subjective evaluation factors.Faster and more reliable decision-making. Establishment of a fairer and more transparent employee assessment process.

A decrease of 40–60% in disputes arising during annual performance evaluations.The study further shows that the combination of NLP techniques, outlier detection algorithms, and semantic vectorization effectively eliminates the key data-quality issues commonly observed in KPI systems (Table 3).

Table 3. Advantages and Limitations

Advantages	Limitations
Full automation of data processing. The system is capable of handling heterogeneous data formats. Incorrect or inconsistent values are detected immediately. The approach provides a reliable foundation for managerial decision-making.	Training the model requires a large dataset. Semantic normalization demands substantial computational power. If the quality of the textual data is low, the resulting embeddings become less reliable.

CONCLUSIONS AND SUGGESTIONS

The proposed AI-driven normalization methodology has demonstrated substantial improvements in the accuracy, consistency, and overall reliability of KPI data processing. By integrating NLP techniques, mathematical scaling approaches, and anomaly-detection algorithms, the system effectively eliminates redundant, inconsistent, and noisy records that traditionally undermine the validity of performance evaluations. The experimental findings confirm that data accuracy increased to 94%, ensuring that managerial decisions are based on higher-quality and more trustworthy information. Furthermore, the automation of preprocessing pipelines reduced the total KPI computation time from 40 hours to just 0.8 hours—a 50-fold improvement in operational efficiency.

Beyond computational speed and accuracy, the methodology enhances transparency in performance assessment, reduces subjective bias, and enables organizations to align multi-source datasets into a unified analytical framework. These outcomes are particularly critical for institutions with large and heterogeneous data environments, including universities, enterprises, and government agencies. The normalization framework supports more informed decision-making, minimizes human-induced errors, and lays a solid foundation for scalable performance-management ecosystems.

Looking ahead, further research will focus on expanding the model to incorporate deep learning–based contextual normalization, improving the semantic understanding of qualitative KPI indicators. Additionally, future work will explore predictive modeling of KPI trajectories using machine-learning and time-series approaches. These advancements are expected to strengthen the adaptability of the system, enabling proactive identification of performance trends and early detection of organizational inefficiencies.

Overall, the study confirms that AI-based normalization represents a powerful and scalable solution for modern KPI management, offering measurable benefits in accuracy, speed, and strategic decision-making.

List of used literature:

1. Shuhratov, Ma'murjon Shuhrat o'g'li, and Jasurbek Olyorbek o'g'li Baxodirov. "An AIDriven Approach to Employee Task and Training Recommendations Using Matrix Factorization." Digital Transformation and Artificial Intelligence: Problems, Innovations and Trends (DTAI-2024): 1st International Scientific-Practical Conference, Tashkent State University of Economics, 11 Sept. 2024, pp. 380–384.
2. Shukhratov, Ma'mur, and Jasur Baxodirov. "Modern Technologies and Methods for Employee Evaluation: Practical Opportunities and Contemporary Challenges." Research Focus International Scientific Journal, vol. 4, no. 5, 2025, pp. 1–15. Research Focus, Uzbekistan. https://doi.org/10.5281/zenodo.15629871

3. Ma'mur Shukhratov, and Jasur Baxodirov,Omonkhonov, Saidkarim, "Системы оценки эффективности работы сотрудников в управлении персоналом на основе искусственного интеллекта." Advances in Science and Humanities, vol. 1, no. 3, 2025, pp. 21–24. https://doi.org:10.70728/human.v01.i03.007
4. Han, Jiawei, Micheline Kamber, and Jian Pei. "Data Preprocessing." Data Mining: Concepts and Techniques, Morgan Kaufmann, 2012, pp. 83–124.
5. Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." Proceedings of NAACL-HLT, 2019, pp. 4171–4186.
6. Reimers, Nils, and Iryna Gurevych. "Sentence-BERT: Sentence Embeddings Using Siamese BERT-Networks." Proceedings of EMNLP/IJCNLP, 2019, pp. 3982–3992. https://doi.org/10.48550/arXiv.1908.10084
7. Kaplan, Robert S., and David P. Norton. "Using the Balanced Scorecard as a Strategic Management System." Harvard Business Review, vol. 74, no. 1, 1996, pp. 75–85.
8. Liu, Fei Tony, Kai Ming Ting, and Zhi-Hua Zhou. "Isolation Forest." 2008 IEEE International Conference on Data Mining, 2008, pp. 413–422. https://doi.org/10.1109/ICDM.2008.17
9. Cascio, Wayne F., and Herman Aguinis. "HR Measurement and Analytics." Applied Psychology in Human Resource Management, Pearson, 2020, pp. 110–145.
10. Chandola, Varun, Arindam Banerjee, and Vipin Kumar. "Anomaly Detection: A Survey." ACM Computing Surveys, vol. 41, no. 3, 2009, pp. 1–58.

Proofreader: Zokir ALIBEKOV

Layout and Designer: Oloviddin Sobir ugli

2025. № 12

To publish articles in this journal, you may submit articles, advertisements, stories, and other creative materials through the following links. Materials and advertisements are published on a paid basis.

You may subscribe to the journal at any time using the following details. Once subscribed, please send a screenshot or photo of your payment confirmation to our Telegram page @iqtisodiyot_77. Based on this, we will send the latest issue of the journal to your address each month.

"The journal "INNOVATION SCIENCE AND TECHNOLOGY" has been registered by the Agency for Information and Mass Communications under the Administration of the President of the Republic of Uzbekistan from 09.10.2024 under the registration number №390637. License number: C-5669633. PNFL: 30407832680027

Our address: Tashkent city, Yunusobod district, 19th block,
House 17.

TASHKENT STATE UNIVERSITY OF ECONOMICS

Acceptance of articles

Published every monthly

Directions

Social, economic, political, technological, scientific

Scopus || Scientific electronic journal specializing in Scopus

CERTIFICATE NUMBER: №390637

ORDER NUMBER ACCORDING TO THE LICENSE REGISTER: C-5669633

CONTACT:

Contact us
+998 50 737 87 88

Telegram channel
t.me/scopus_IST2100

Journal official website
https://ist-journal.uz/index.php/IST