Shahrood University of Technology

Journal of Mining and Environment (JME)

Journal homepage: www.jme.shahroodut.ac.ir

Iranian Society of Mining Engineering (IRSME)

Productivity Evaluation in Open Pit Mining Using Machine Learning Methods

Heydar Bagloo[1*], and Mohsen Soleiman Dehkordi[2]

1. Research and Development Unit, Dispatching center, Chadormalu Mining Complex, Yazd, Iran
2. Management Department of Chadormalu Mining Complex, Yazd, Iran

Article Info

Received 1 February 2025

Received in Revised form 5 March 2025

Accepted 28 April 2025

Published online 28 April 2025

DOI: 10.22044/jme.2025.15692.3017

Keywords

Open-pit mining

Truck

Shovel

Machine learning

Supervised learning

Abstract

Loading and haulage operations in open-pit mining represent a significant portion of overall costs. Among various load and transport systems, the shovel-truck method is favored for its flexibility. Consequently, extensive research has been conducted to optimize this system, resulting in numerous productivity-enhancing methods. However, evaluating the effectiveness of these optimization techniques, particularly in short-term mining activities under varying operational conditions, remains essential. Additionally, understanding how changes in operational conditions impact productivity is important for addressing production fluctuations in daily mining operations. To tackle these challenges, this study uniquely applies advanced machine learning techniques to short-term mining planning, resulting in the development of a real-time Productivity Evaluation Model (PEM) based on supervised learning methods for optimizing truck-shovel operations in open-pit mining. The model, developed and tested using data from a large-scale mining operation in Iran, demonstrated that the Decision Tree was the most effective, achieving an R^2 value of 0.96. This was closely followed by Random Forest and Gradient Boosting, both with R^2 values of 0.95. However, the choice of the most suitable learning method may vary depending on the specific dataset and context. The model determines the most appropriate learning method for each dataset within specific mining operations.

1. Introduction

As the global population and technological advancements continue to rise, the demand for mining products increases accordingly [1]. However, the costs of mining operations are also increasing significantly [2], and environmental regulations are becoming stricter [3]. Consequently, mining operations must optimize their processes to reduce costs and align with sustainable development objectives. Likewise, open-pit mining, as the most common method of surface mining with high cost expenditures, has been the focus of numerous optimization studies in recent years [4, 5].

In open-pit mining, the main operations include drilling, blasting, loading, and haulage. Among these steps, loading and haulage operations are particularly important. They account for more than 50 percent of total mining costs [6-8]. Additionally, the effectiveness of this stage is a key factor in shaping the production rate that the mining operation can achieve [9]. Among the various loading and transport methods, the truck-shovel system remains the primary approach in open-pit mining operation [10], due to their flexibility [11], advancements in technology, and the increasingly challenging conditions of mining operations caused by natural and technological factors [10]. However, it is widely recognized that truck-shovel systems require a substantial investment [11-13].

In general, numerous studies in the literature on reducing mining expenses, decreasing energy consumption, and mitigating the environmental impact of truck-shovel systems demonstrate that

Corresponding author: hbagloo@gmail.com

continuous optimization of this method is crucial in mining operations. There is a substantial amount of literature in this area; however, we present a selection of examples to illustrate the research conducted in this field. The work of Munirathinam and Yingling provides a detailed classification of older truck dispatching strategies, offering analyses of their mathematical formulations, alongside an evaluation of their strengths, weaknesses, and expected performance in practical applications [14]. Souza et al. address the Open-Pit-Mining Operational Planning problem with dynamic truck allocation, proposing a hybrid algorithm combining Greedy Randomized Adaptive Search Procedures and General Variable Neighborhood Search [15]. Their computational experiments demonstrate the algorithm's efficiency in finding near-optimal solutions with minimal computing times, validated against CPLEX optimizer results. Ta et al. develop models to minimize the number of trucks required for a specific set of shovels, utilizing the nonlinear relationship between a shovel’s idle probability and the number of trucks allocated to it [16]. Zhang and Xia employ an integer programming approach to develop a method for determining the number of truck trips between loading and dump sites [17]. Mohtasham et al. develop a chance-constrained goal programming (CCGP) model to optimize truck allocation in open-pit mines, achieving over 20.6% improvement in production rate compared to traditional methods. Their work demonstrates the effectiveness of integrating advanced optimization techniques, such as CPLEX software, in mining operations to meet multiple objectives, including cost reduction and maintaining desired grade and stripping ratios [18]. This approach demonstrates a reduction in truck operating costs by 15.65% compared to a fixed truck assignment policy for a homogeneous fleet. Patterson et al. introduce a mixed integer linear programming method to optimize haulage scheduling and reduce energy consumption for trucks and shovels, validating their approach with a case study from a mine in South East Queensland [19]. Afrapoli et al. create a multi-objective transportation model for real-time truck dispatching aimed at minimizing shovel idle times, truck wait times, and deviations from production targets [20]. The findings indicate that the model effectively meets production requirements using a fleet 15% smaller than the calculated size and achieves full processing plant capacity with 30% fewer trucks. Ghaziani et al. optimize the transport fleet of the Sungun copper mine by comparing flexible dispatching with fixed allocation methods and employing the firefly metaheuristic algorithm to find the most suitable transportation arrangement, resulting in significant productivity and cost improvements [21]. Their study demonstrates the effectiveness of flexible dispatching and heterogeneous truck fleets, achieving substantial increases in production rate and reductions in idle time and operating costs. Mirzaei-Nasirabad et al. introduce a method for optimizing truck-shovel systems in open-pit mining using a multi-stage dispatching approach [22]. Initially, a scenario-based method is proposed to identify the optimal truck size for the operation. Subsequently, a multi-objective mathematical model for dynamic truck allocation is presented, focusing on reducing fleet waiting times and minimizing deviations from production targets.

While the optimization methods derived from these studies generally prove beneficial and improve productivity in open pit mining, it is essential to have methodologies to measure and compare their effectiveness for informed decision-making. The parameters often used in the literature for optimizing truck-shovel operations include shovel idle times, truck waiting times, deviations from production requirements, truck and shovel energy consumption, the number of trucks for a given set of shovels, and overall operating costs. However, comparing optimization models based on these factors—whether through simulation or against previous methods—can introduce errors due to the continuously changing operational conditions in a mine. Variables such as the distance between loading and unloading areas, path slopes and quality, fleet capacities and lifespans, and the stripping ratio can all fluctuate over time. As a result, applying an optimization method and comparing its outcomes with past operations may not yield accurate results.

In this paper, we introduce a Productivity Evaluation Model (PEM) that employs supervised machine learning techniques to assess operational efficiency under real-time conditions. The model utilizes data related to mining operational conditions to generate an hourly production figure, enabling the evaluation of production outcomes derived from optimization methods and determining whether they meet or fall short of the required target under specific operational conditions. Additionally, it facilitates the assessment of daily production by comparing it to the required target for the given conditions on a particular day of operation. The model is designed to be flexible, allowing parameter adjustments based on industry expertise while supporting the integration of various machine learning techniques. To enhance its robustness, future research could explore other learning methods and feature engineering to improve accuracy, while external validation using data from other mining operations would help assess adaptability. Additionally, sensitivity analysis could provide insights into how variations in input features impact model predictions, enhancing reliability. The remainder of this paper is structured as follows: Section 2 outlines the methodology, Section 3 provides information about the case study, Section 4 presents the results, and Section 5 concludes the study with discussions and suggestions for future research.

2. Methodology

In this study, we aim to develop a Productivity Evaluation Model (PEM) to predict production based on real-time operational conditions, including haulage distance (the cumulative total distance traveled by all trucks in a full cycle, encompassing both the loaded haulage and empty return), stripping ratio, and the number of trucks available per shovel. These features were selected because they are directly related to the production rate. Haulage distance impacts hauling efficiency, stripping ratio determines the material moved per cycle, and the number of trucks per loader ensures optimal fleet utilization. To achieve this, we collected daily data over a seven-month period from a large-scale mining area in Iran, covering both cold and warm seasons, thereby implicitly accounting for seasonal effects. The data includes haulage distance (daily kilometers per service), hauled ore and waste (daily tonnage), the number of trucks available per shovel (based on loading and hauled tonnage for the mine's loading and haulage fleet), and daily production (services per hour). While other potential features, such as weather conditions, truck maintenance schedules, and operator efficiency, were not included in this study, they could be considered in future research to improve the model's accuracy.

The collected data were preprocessed and cleansed to ensure consistency. First, correlation analysis was conducted for feature selection to retain the most relevant variables while minimizing redundancy. Then, missing values were replaced with the mean of the respective features. Outliers, defined as values that deviate significantly from the central tendency of the data, were replaced with the nearest valid observation to maintain data integrity. Following this, normalization was applied using the Min-Max method to scale the data within a consistent range.

Subsequently, a set of supervised learning methods was employed to develop predictive models, aiming to analyze and improve the accuracy of real-time production estimation. The primary objective was to design a model capable of integrating additional methods seamlessly, allowing for future experimentation with alternative approaches. The selection of the five specific methods was based on their distinct advantages in handling different data structures, ensuring a balanced comparison of predictive performance. Given the relatively small dataset, methods that perform well with limited data were prioritized.

Linear Regression (LR) was chosen as a baseline due to its simplicity and interpretability, providing a fundamental benchmark for model evaluation. Decision Trees (DT) and Random Forest (RF) were included to capture non-linear relationships and demonstrate the benefits of ensemble learning in reducing variance. These tree-based models are also well-suited for small datasets, as they do not rely on large amounts of training data to generalize effectively. Gradient Boosting (GB) was selected for its iterative approach to minimizing errors and enhancing prediction accuracy, particularly useful when working with limited data. Finally, Support Vector Regression (SVR) was incorporated for its ability to model complex patterns using kernel-based transformations, which can be advantageous in small datasets where high-dimensional relationships need to be captured.

We then evaluated the performance of these models using two key metrics: the coefficient of determination (R^2) and mean squared error (MSE). While both metrics were considered, the primary focus in our comparisons is on R^2 due to its direct relevance to the explanatory power of the models. The models were implemented in Python using libraries including NumPy, Pandas, and Scikit-learn. Data visualization was conducted through Matplotlib and Seaborn to effectively present and analyse the results.

3. Case study

The productivity evaluation model presented in this paper utilizes data from a specific mining operation; however, it is designed to be adaptable to other mining contexts. The model allows for the integration of data from any mining operation, enabling the assessment of performance and the optimization of operations based on the specific dataset provided. This flexibility enables the model to be applied across different mining sites, as long as relevant and comparable data is available.

The model was developed using data from a large-scale open-pit mining operation in Iran. In this operation, large shovels load ore and waste material into haul trucks, which then transport the material to designated processing or stockpiling areas. While the efficiency of the truck-shovel system depends on several operational factors, we specifically selected haulage distance, stripping ratio, and the number of trucks per loader due to their direct impact on the model's target (hourly production).

Haulage distance is calculated as the cumulative distance traveled by all trucks divided by the total number of haulage services performed. This metric reflects the average transportation effort required per cycle, influencing fuel consumption, cycle times, and overall productivity. The number of trucks per loader is determined by the total truck working hours in a day divided by the total shovel working hours in that same period. This ratio helps assess whether the fleet allocation is balanced, ensuring that neither shovels nor trucks experience excessive idle times.

Instead of directly using the stripping ratio, which could be undefined when ore tonnage is zero on certain days for some contractors, the model considers both ore and waste quantities. This adjustment ensures a stable representation of material movement, providing a more reliable metric for evaluating mining efficiency.

4. Results

The results section is structured into two distinct parts to provide a comprehensive analysis. In the first part, we examine and describe the dataset, focusing on the distribution of key features and their interrelationships. The second part then addresses the evaluation of our machine learning models, emphasizing their predictive accuracy and suitability for real-time mining operations.

4.1. Descriptive Analysis and Data Exploration

The analysis of the collected data produced several significant outcomes that provide valuable insights into the effectiveness of the applied methodologies. As an initial step, we conducted a comprehensive investigation of the data. The description of the data, including mean, standard deviation (std), minimum and maximum values, and skewness is presented in Table 1. According this Table, the average haulage distance is about 7 km per service, with waste and ore quantities around 32,000 tons and 4,700 tons, respectively. The typical number of trucks per loader is around 4, and the production rate averages about 2.3 services per hour. The skewness values indicate that haulage distance (-0.08) and waste quantities (-0.18) are nearly symmetric, with no significant tails. Ore material (1.16) and production rate (0.91) are positively skewed, showing a long tail on the right, suggesting the presence of higher extreme values. Truck per loader (0.31) has a mild positive skewness, indicating a slight right tail.

Table 1. Descriptive Statistics of Parameters

Statistics	Distance (km/ser)	Waste (ton)	Ore (ton)	Truck per loader	Production (ser/h)
Mean	7.08	32061.39	4761.11	3.66	2.35
Std	1.67	9194.76	5448.13	0.6	0.47
Min	3.6	6900	0	2	1.64
Max	11.8	54875	26440	6.5	3.61
Skewness	-0.08	-0.18	1.16	0.31	0.91

Figure 1 provides a graphical representation of the distribution for each model feature through histograms. As shown in the figure, the haulage distance and waste material weight exhibit a relatively even distribution around their respective means. In contrast, the ore material weight predominantly clusters at lower values, specifically below 10,000 tons. Additionally, the number of trucks tends to be concentrated within the range of approximately 2.5 to 4.5.

The relationship between each feature and the target variable was examined. Figure 2 depicts the relationship between production and both haulage distance and the number of trucks per loader. The figure shows that increasing both haulage distance and the number of trucks per loader tends to decrease hourly production. Specifically, although increasing the number of trucks can enhance total daily production, it also leads to longer waiting times at loading points, which reduces hourly production. Whereas both factors contribute to a decrease in hourly production, haulage distance has a greater effect on production compared to the number of trucks.

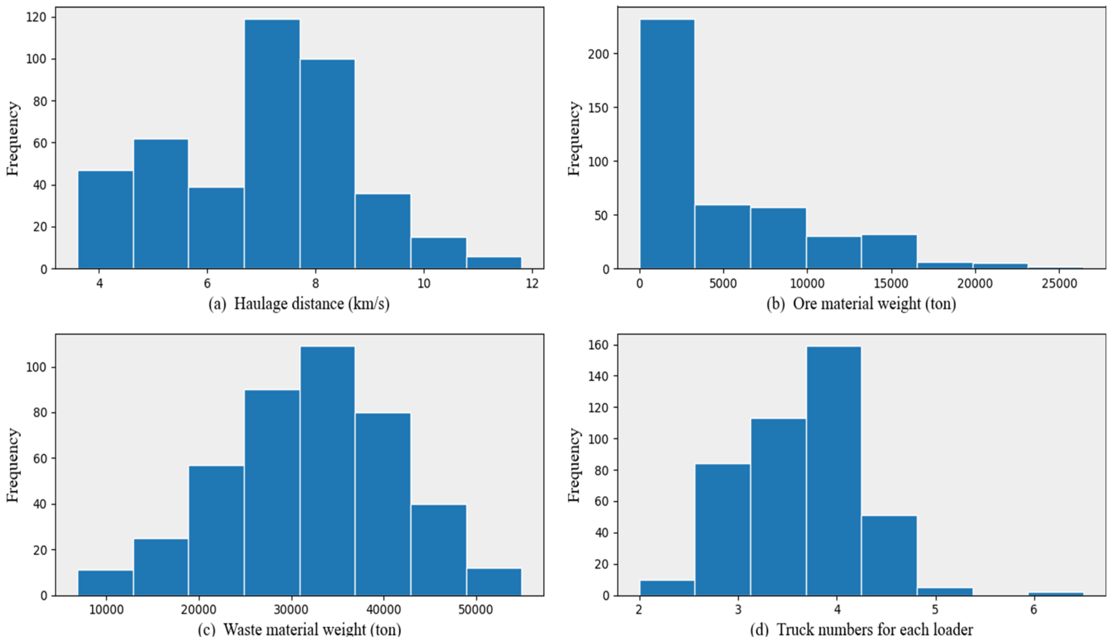

Figure 1. Distribution of model features

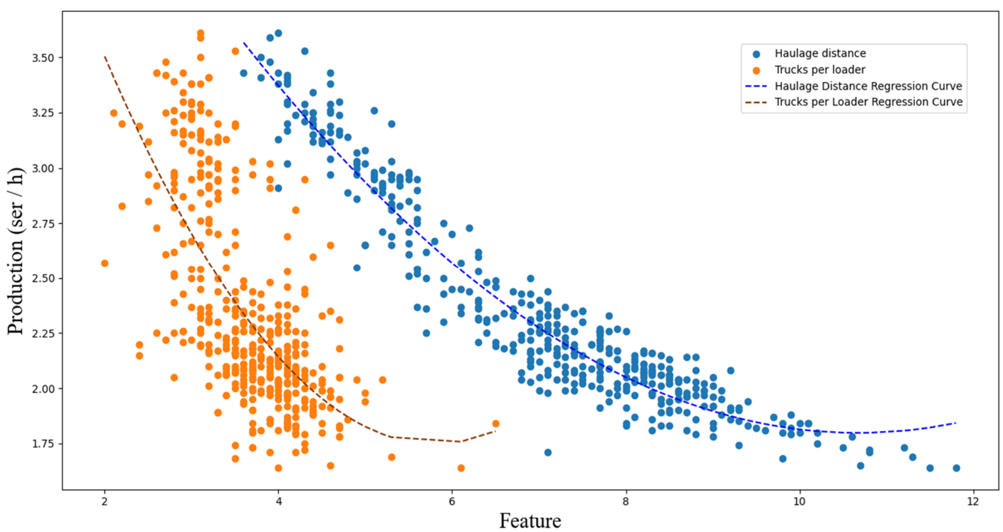

Figure 2. Effects of haulage distance and number of trucks assigned per loader on hourly production.

Figure 3 illustrates the effects of hauled ore and waste material weights on hourly production. While the weight of ore has a minimal impact on production, the weight of waste material has a more pronounced effect on daily production. This may be attributed to the fact that waste material in the studied mine is primarily loaded and transported from higher elevations, which results in shorter haulage distances and, consequently, higher hourly production.

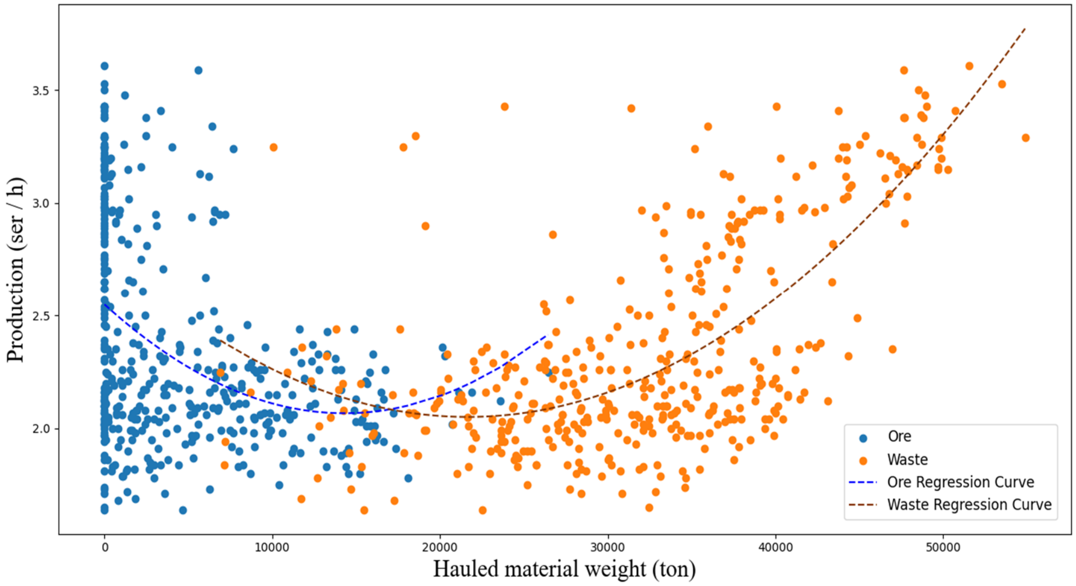

Figure 3. Effects of hauled ore and waste material weights on hourly production.

The relationships between the model features and the target variable were examined to understand their dependencies and impact on production. Figure 4 presents a correlation matrix depicting the relationships both among features and between features and hourly production. The correlation coefficients are as follows:

- Haulage distance has a strong negative correlation of -0.93 with hourly production, indicating that longer haulage distances are strongly associated with reduced hourly production.
- Waste weight has a moderate positive correlation of 0.59 with hourly production, meaning that as the weight of waste material increases, hourly production tends to increase as well, though the relationship is not as strong as that with haulage distance.
- Ore weight has a weaker negative correlation of -0.35 with hourly production, suggesting that higher ore weight is somewhat linked to lower production, but the effect is relatively minor. This may be attributed to the longer distances between ore extraction sites and unloading areas compared to those for waste. Increased ore weight often implies longer haulage distances, which can lead to reduced hourly production.
- The number of trucks assigned to each loader shows a moderate negative correlation of -0.64 with hourly production, indicating that an increase in trucks per loader is generally associated with lower hourly production. As previously explained, more trucks per loader lead to longer waiting times, increased haul cycle times, and ultimately reduced hourly production.

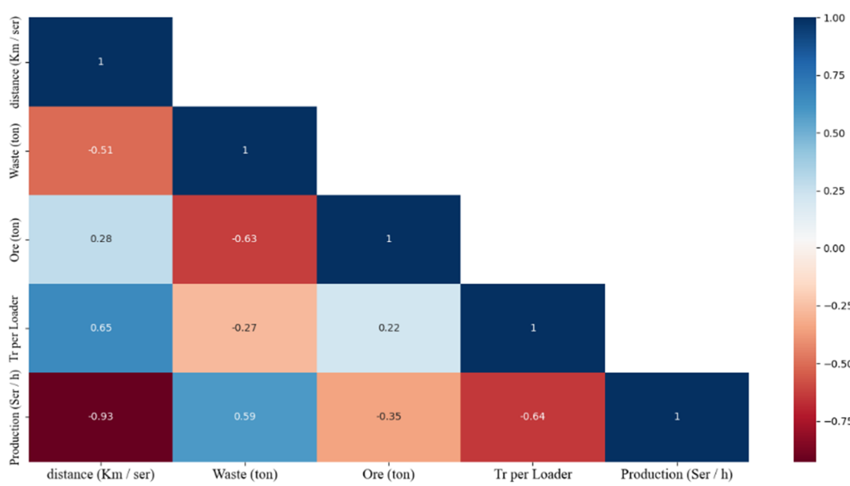

Figure 4. Correlation matrix depicting relationships among features and hourly production.

4.2. Modeling and Performance Evaluation

After analyzing the data, we partitioned it into training and testing sets, reserving 20% for testing and using the remaining data to train the model. To achieve this, we applied supervised learning techniques, specifically Linear Regression (LR), Decision Tree (DT), Random Forest (RF), Gradient Boosting (GB), and Support Vector Regression (SVR). The performance of these models was evaluated using the coefficient of determination (R^2), which quantifies the proportion of variance in the target variable explained by the model. The primary goal of this study is to develop a predictive model tailored for real-time mining operations, given their dynamic and complex nature. While the current framework focuses on these specific algorithms, the proposed model is designed to be adaptable, allowing for the integration and evaluation of additional learning techniques in future research to further enhance performance and applicability.

Initially, the Linear Regression (LR) model demonstrated a coefficient of determination (R^2) value of 0.90, reflecting its effectiveness in explaining the variance within the dataset. We then explored the potential benefits of regularization techniques, including Ridge and Lasso regression, to determine if they could enhance the model's performance further. However, these methods did not improve the performance beyond that of the standard LR model. The R^2 values obtained with regularization were lower than those achieved with the basic LR model, indicating that regularization did not enhance the model's effectiveness for this particular dataset and feature set. This suggests that regularization may not be necessary for our data, where the risk of overfitting is relatively low.

After evaluating the Linear Regression model, we proceeded to assess the Decision Tree (DT) model, which initially achieved an R^2 value of 0.92. However, by fine-tuning key parameters—specifically setting the maximum depth to 5, the minimum samples required to split a node to 12, and the minimum samples per leaf to 3 (see Table 2) —the R^2 value improved significantly to 0.96. In the fine-tuning process, we manually tested different combinations of key hyperparameters, rather than using methods like Grid Search or Random Search, as we aimed for a more focused, domain-driven approach. However, these methods could be considered in future projects to assess their potential for further optimization.

Table 2. Fine-Tuned Hyperparameters for the Decision Tree (DT) Method and R^2 Values

Hyperparameter	Default Value	Final Value
Maximum Depth	None (unlimited)	5
Min Samples Split	2	12
Min Samples Leaf	1	3
R^2 value	0.92	0.96

The Random Forest model was configured with 100 trees, and a minimum of 2 samples required to split an internal node. The Gradient Boosting model used 100 estimators, a learning rate of 0.1, and a maximum depth of 3. Both models achieved an R^2 value of 0.95. Despite testing various parameters no significant improvement in the R^2 value was observed.

Finally, we applied the Support Vector Regression (SVR) model with a radial basis function (RBF) kernel, a regularization parameter of 1.0, and a tolerance of 0.001. This model achieved an R^2 value of 0.93. Despite experimenting with different settings for the regularization parameter, various kernel types, and different gamma values for the Radial Basis Function (RBF) kernel, no significant improvement in the R^2 value was observed.

Figure 5 displays the R^2 values for various methods, with each plot comparing predicted values to actual values for the corresponding method. As presented in Figure 5 and Table 3, the Decision Tree model achieved the highest R^2 value of 0.96. This was followed by the Random Forest and Gradient Boosting methods, both with R^2 values of 0.95. Linear Regression had an R^2 value of 0.90, while the Support Vector Regression method achieved an R^2 value of 0.93.

To further evaluate the performance of the models, we analyzed the residuals to gain insight into the discrepancies between predicted and actual values. Figure 6 presents the residual plots for each model, highlighting the distribution and spread of residuals.

- **Decision Tree (DT), Random Forest (RF), and Gradient Boosting (GB) Models:** The residuals are randomly distributed around the zero line, indicating no discernible pattern. This suggests that these models effectively capture the underlying data trends and do not exhibit noticeable bias. Visual inspection of the residual plots suggests that the residuals are approximately normally distributed, which further supports the robustness of these models.

- **Linear Regression (LR) Model:** The residuals for data points with target values between 2 and 2.8 (normalized values 0.2 and 0.6) are predominantly below the zero line, while residuals for data points with target values outside this range tend to be above the zero line. This indicates that the LR model may exhibit bias in its predictions, particularly for values outside the central range. Additionally, the residuals do not appear to be normally distributed, as there is a noticeable deviation from normality and non-constant variance.

- **Support Vector Regression (SVR) Model:** The residuals for the Support Vector Regression (SVR) model are concentrated around the zero line, suggesting minimal systematic bias. This indicates that the SVR model is performing effectively for this dataset.

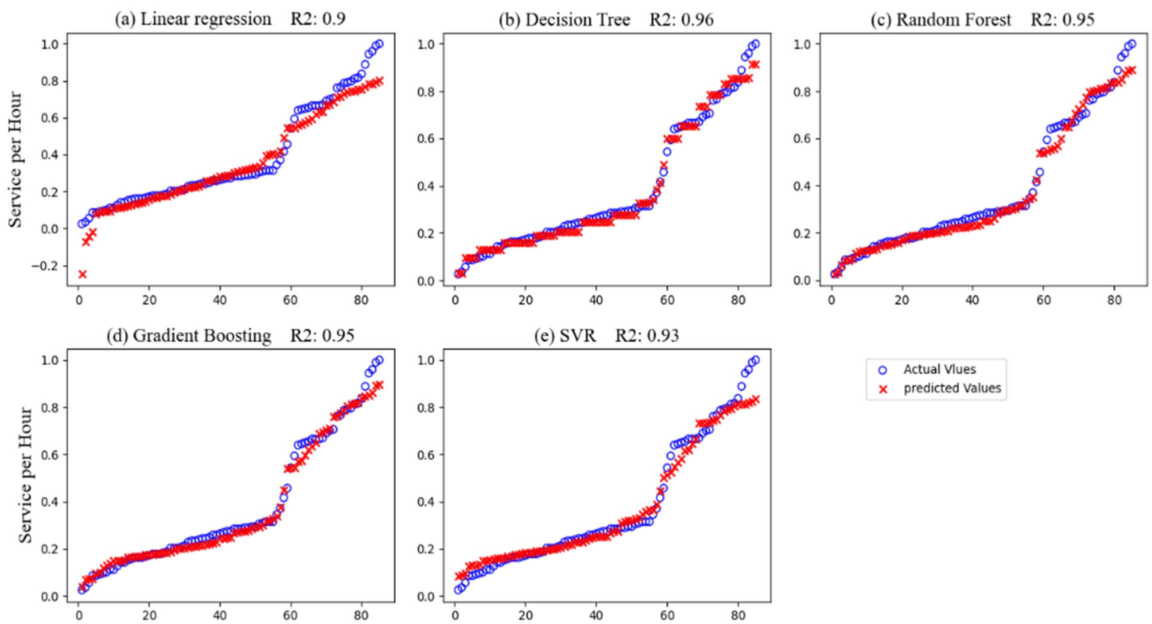

Figure 5. R² values and comparison plots of predicted versus actual values for different learning methods.

Table 3. Model Performance Metrics

Performance Metrics	LR	DT	RF	GB	SVR
Mean Squared Error (MSE)	0.007	0.003	0.004	0.004	0.005
Coefficient of determination (R2)	0.90	0.96	0.95	0.95	0.93

To investigate the distribution and spread of residuals and provide a visual summary of each model's prediction accuracy, the box plots of residuals are presented in Figure 7. Based on this figure, the following observations can be made:

- **Linear Regression:** Exhibits the largest spread in residuals, indicating higher variance in predictions. The presence of outliers suggests instances of large prediction errors, and the median residual is slightly above zero.

- **Decision Tree:** Displays a relatively compact interquartile range, indicating consistent predictions. However, outliers are present, reflecting occasional large errors. The median residual is close to zero.

- **Random Forest:** Shows a balanced residual distribution with a compact interquartile range, suggesting stable predictions. The median residual is close to zero, and fewer extreme outliers are present compared to other models.

- **Gradient Boosting:** Demonstrates a stable residual distribution with low variability. The median residual is near zero, and the spread of residuals is relatively small, indicating strong predictive performance.

- **Support Vector Regression (SVR):** Exhibits a moderate spread in residuals, with the median residual very close to zero. Outliers are present, but the overall distribution suggests a reasonable level of consistency in predictions.

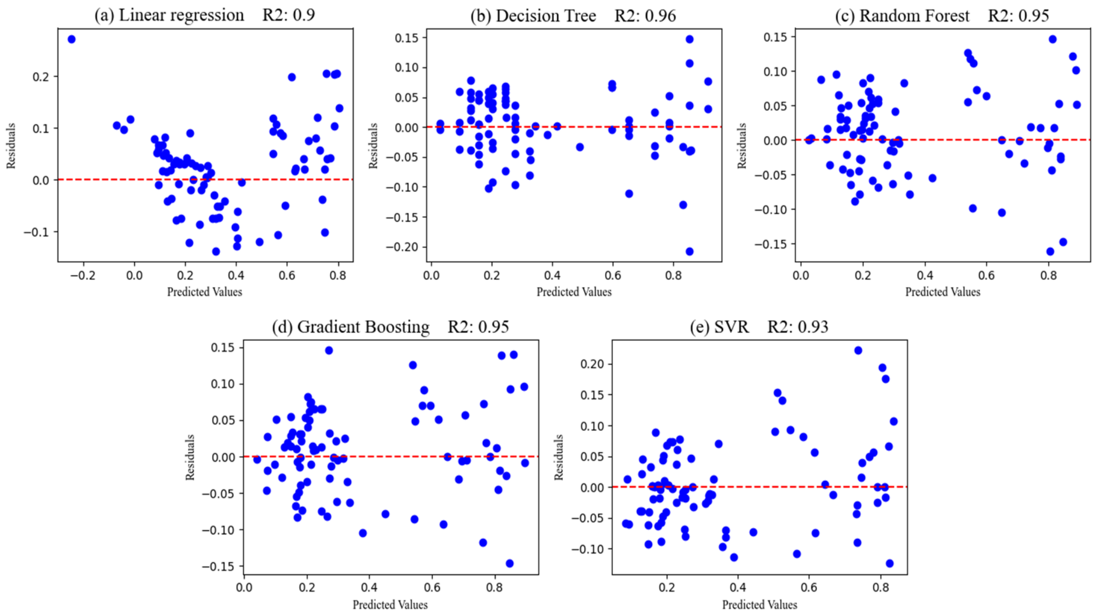

Figure 6. Residual plots for applied models, showing the distribution of residuals and patterns in predictions.

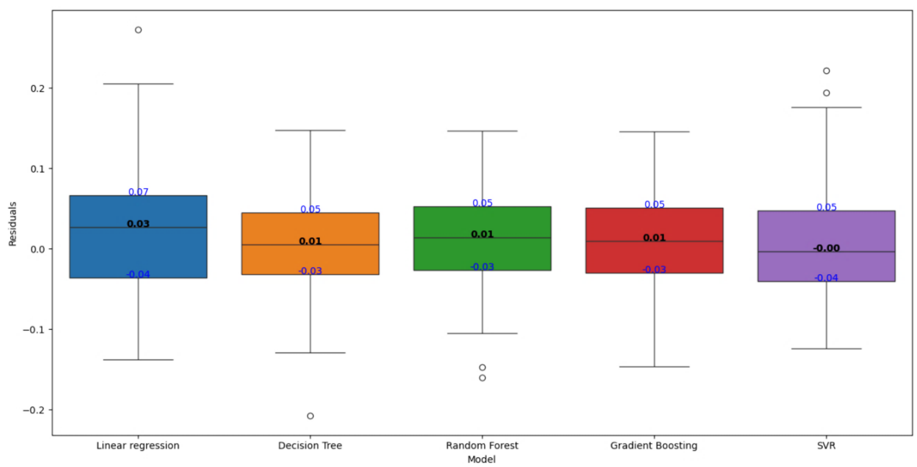

Figure 7. Boxplots of residuals for applied learning methods

Finally, the Taylor diagram was used to evaluate the models' performance in terms of their correlation, standard deviation, and mean squared error, as shown in Figure 8. The diagram reveals that the Decision Tree, Random Forest, and Gradient Boosting models exhibit strong performance, clustering closely with high correlation and low mean squared error. Linear Regression also demonstrates acceptable performance with high correlation and low MSE, though its standard deviation is slightly different. Support Vector Regression (SVR) is positioned similarly to the other models, showing a high correlation and low standard deviation, indicating that its performance is comparable to the best-performing models.

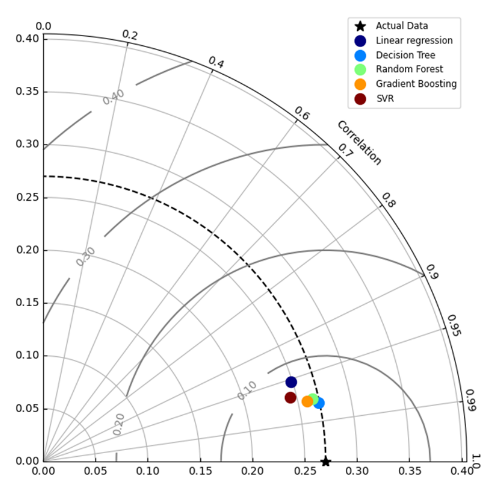

Figure 8. Performance comparison of applied learning methods using Taylor diagram

5. Conclusions

Truck-shovel systems are fundamental to open-pit mining, with the ability to significantly influence operational efficiency and costs. While optimization of these systems has been extensively studied, evaluating their performance under dynamic conditions remains challenging. Current research has focused on reducing operational costs, energy consumption, and environmental impact, but there is a gap in methodologies that effectively assess and compare the real-time productivity outcomes of different operational strategies. This paper addresses this shortcoming by introducing a Productivity Evaluation Model (PEM) that uses supervised learning techniques to predict and compare the productivity of truck-shovel systems under varying conditions. This contribution provides valuable insights into improving truck-shovel efficiency and lays the groundwork for future research to refine these methods and expand their applicability.

The performance of the model discussed in this study may vary across different mining sites, depending on the specific input data. To address this, our model is designed to incorporate data from each mine individually and selects the most suitable machine learning method based on the characteristics of that particular dataset. This enables the model to be applied to other mining operations. However, a notable limitation is the model's sensitivity to the quality and quantity of the available data, which can impact its overall performance. Therefore, to better understand the model's adaptability and reliability across diverse mining environments, further validation with datasets from multiple mines would be essential.

The model was developed using data from a large-scale mining operation in Iran. Supervised learning methods, including Linear Regression, Decision Tree, Random Forest, Gradient Boosting, and Support Vector Regression, were employed to predict hourly production. Key features included haulage distance, material weights, and truck availability per shovel. Model performance was assessed using R^2 and mean squared error (MSE) to ensure predictive accuracy.

The Decision Tree model, after fine-tuning key parameters such as maximum depth, minimum samples to split, and minimum samples per leaf, achieved the highest R^2 value of 0.96. This was followed closely by the RF and GB models, both with an R^2 value of 0.95. These models demonstrated strong performance, clustering closely in the Taylor diagram with high correlation and low MSE. SVR had a high R^2 value of 0.93, indicating strong predictive performance. Finally, Linear Regression (LR) had the lowest R^2 value (0.90), indicating that it may not be as effective as the other models in predicting the target variable for this dataset.

Overall, Decision Tree, Random Forest, and Gradient Boosting models demonstrated the highest effectiveness for the dataset. SVR performed well, while Linear Regression showed acceptable performance but was not as effective as the other models for the given dataset. Future research could explore additional learning methods and feature engineering to further enhance accuracy.

References

[1]. Ali, S., et al. (2017). Mineral supply for sustainable development requires resource governance. *Nature, 543*(7645), 367-372.

[2]. Humphreys, D. (2020). Mining productivity and the fourth industrial revolution. *Mineral Economics, 33*(1), 115-125.

[3]. Kinnunen, P.H.M., & Kaksonen, A.H. (2019). Towards circular economy in mining: Opportunities and bottlenecks for tailings valorization. *Journal of Cleaner Production, 228,* 153-160.

[4]. Blom, M., Pearce, A., & Stuckey, P. (2018). Short-Term Planning for Open Pit Mines: A Review. *International Journal of Mining, Reclamation and Environment, 33*.

[5]. Moradi, A., & Askari Nasab, H. (2015). Open-Pit Mine Production Optimization: A Review of Models and Algorithms. *Engineering,* 251436017.

[6]. Gamache, M., & Alarie, S. (2002). Overview of Solution Strategies Used in Truck Dispatching Systems for Open Pit Mines. *International Journal of Surface Mining, Reclamation and Environment, 16,* 59-76.

[7]. Elijah, K., et al. (2021). Optimisation of shovel-truck haulage system in an open pit using queuing approach. *Arabian Journal of Geosciences, 14*(11), 973.

[8]. Cerna, G.P., & Obredor-Baldovino, T. (2024). A Mixed Integer Programming optimization model for mining truck dispatch policies using traffic constraints: Case of a copper mine in northern Chile. *Procedia Computer Science, 231,* 660-665.

[9]. Aguayo, I.A., Nehring, M., & Ullah, G.M.W. (2021). Optimising Productivity and Safety of the Open Pit Loading and Haulage System with a Surge Loader. *Mining, 1,* 167-179.

[10]. Kuznetsov, D., & Kosolapov, A. (2022). Dynamic of performance of open-pit dump trucks in ore mining in severe climatic environment. *Transportation Research Procedia, 63,* 1042-1048.

[11]. Bakhtavar, E., & Mahmoudi, H. (2020). Development of a scenario-based robust model for the optimal truck-shovel allocation in open-pit mining. *Computers & Operations Research, 115,* 104539.

[12]. Bajany, D.M., Xia, X., & Zhang, L. (2017). A MILP Model for Truck-shovel Scheduling to Minimize Fuel Consumption. *Energy Procedia, 105,* 2739-2745.

[13]. Makarova, I., et al. (2023). Improving the Environmental Friendliness of the Mining Complex Through Alternative Fuel for Mine Dump Trucks. *Transportation Research Procedia, 68,* 755-760.

[14]. Munirathinam, M., & Yingling, J.C. (1994). A review of computer-based truck dispatching strategies for surface mining operations. *International Journal of Surface Mining, Reclamation and Environment, 8*(1), 1-15.

[15]. Souza, M.J.F., et al. (2010). A hybrid heuristic algorithm for the open-pit-mining operational planning problem. *European Journal of Operational Research, 207*(2), 1041-1051.

[16]. Ta, C.H., Ingolfsson, A., & Doucette, J. (2013). A linear model for surface mining haul truck allocation incorporating shovel idle probabilities. *European Journal of Operational Research, 231*(3), 770-778.

[17]. Zhang, L. & Xia, X. (2015). An Integer Programming Approach for Truck-Shovel Dispatching Problem in Open-Pit Mines. *Energy Procedia, 75,* 1779-1784.

[18]. Mohtasham, M., Mirzaei Nasirabad, H., & Mahmoodi Markid, A. (2017). Development of a goal programming model for optimization of truck allocation in open pit mines. *Journal of Mining and Environment, 8*(3), 359-371.

[19]. Patterson, S.R., Kozan, E., & Hyland, P. (2017). Energy efficient scheduling of open-pit coal mine trucks. *European Journal of Operational Research, 262*(2), 759-770.

[20]. Moradi Afrapoli, A., Tabesh, M., & Askari-Nasab, H. (2019). A multiple objective transportation problem approach to dynamic truck dispatching in surface mines. *European Journal of Operational Research, 276*(1), 331-342.

[21]. Ghaziania, H.h., et al. (2021). Design of Loading and Transportation Fleet in Open-Pit Mines using Simulation Approach and Metaheuristic Algorithms. *Journal of Mining and Environment, 12*(4), 1177-1188.

[22]. Mirzaei-Nasirabad, H., et al. (2023). An optimization model for the real-time truck dispatching problem in open-pit mining operations. *Optimization and Engineering, 24*(4), 2449-2473.

نشریه مهندسی معدن و محیط زیست
انجمن مهندسی معدن ایران
دانشگاه صنعتی شاهرود
نشانی نشریه: www.jme.shahroodut.ac.ir

ارزیابی بهره‌وری در معادن روباز با استفاده از روش‌های یادگیری ماشین

حیدر بگلو[1*] و محسن سلیمان دهکردی[2]

۱. واحد تحقیق و توسعه، مرکز دیسپچینگ، مجتمع معدنی چادرملو، یزد، ایران
۲. بخش مدیریت مجتمع معدنی چادرملو، یزد، ایران

اطلاعات مقاله

تاریخ ارسال: ۲۰۲۵/۰۲/۰۱
تاریخ داوری: ۲۰۲۵/۰۳/۰۵
تاریخ پذیرش: ۲۰۲۵/۰۴/۲۸

DOI:10.22044/jme.2025.15692.3017

کلمات کلیدی

معدنکاری روباز
کامیون
شاول
یادگیری ماشین
یادگیری نظارت‌شده

چکیده

عملیات بارگیری و حمل در معادن روباز سهم قابل توجهی از کل هزینه‌ها را به خود اختصاص می‌دهد. در میان سیستم‌های مختلف بارگیری و حمل، روش شاول-کامیون به دلیل انعطاف‌پذیری از محبوبیت خاصی برخوردار است. در نتیجه، تحقیقات گسترده‌ای برای بهینه‌سازی این سیستم انجام گرفته که منجر به ابداع روش‌های متعددی برای افزایش بهره‌وری شده است. با این حال، ارزیابی اثربخشی این تکنیک‌های بهینه‌سازی، به ویژه در برنامه های کوتاه‌مدت فعالیت‌های معدنی تحت شرایط عملیاتی متغیر، همچنان از اهمیت بالایی برخوردار است. علاوه بر این، درک تأثیر تغییرات شرایط عملیاتی بر بهره‌وری برای مدیریت نوسانات تولید در عملیات روزانه معدنکاری حیاتی است. برای مواجهه با این چالش‌ها، در این مطالعه با بهره¬گیری از تکنیک‌های پیشرفته یادگیری ماشین در برنامه‌ریزی کوتاه‌مدت معدن، یک مدل ارزیابی بهره‌وری (PEM) مبتنی بر روش‌های یادگیری نظارت‌شده برای بهینه‌سازی عملیات سیستم شاول-کامیون در معادن روباز توسعه داده شده است. برای ایجاد و همچنین ارزیابی عملکرد این مدل از داده‌های عملیاتی یکی از معادن بزرگ مقیاس در ایران استفاده شده است. بر اساس نتایج بدست آمده، درخت تصمیم (Decision Tree) با دستیابی به مقدار ضریب R^2 برابر با ۰.۹۶، مؤثرترین الگوریتم بوده است. الگوریتم¬های جنگل تصادفی (Random Forest) و گرادیان بوستینگ (Gradient Boosting) نیز با مقدار R^2 برابر با ۰.۹۵ عملکرد بسیار نزدیکی به روش درخت تصمیم داشتند. البته، انتخاب مناسب‌ترین روش یادگیری می‌تواند بسته به داده‌ها و شرایط خاص هر معدن متفاوت باشد. با استفاده از این مدل می¬توان بهترین الگوریتم یادگیری متناسب با داده‌های هر عملیات معدنی را انتخاب نمود.

نویسنده مسئول مکاتبات: *hbagloo@gmail.com* (ح. بگلو)