"Research Article"

10.71737/jpm.2026.1192591

Investigating machine learning-based methods to improve maintenance and repair management in oil pipelines with emphasis on weighted KNN

Mohammad Reza Zamani[1], Ahmad Ebrahimi[2]*, Alireza Rashidi Komijan[3]

Abstract

Maintenance costs usually account for more than a third of a manufacturing company's operating costs, so an optimal maintenance strategy and model should improve equipment performance. Therefore, predicting the causes of failures and their roots is a major concern for managers in various industries. In fact, maintenance managers are always looking to discover what factors can lead to failures in production and service continuity in order to prevent them from occurring. This can be even more important in process industries such as oil and gas due to the high volume of losses and losses caused by equipment failures and stoppages in production and product transportation. Predicting pipeline failures and data imbalances is one of the main challenges in the oil industry. Traditional models are unable to accurately identify failures. The nearest neighbor algorithm, as one of the machine learning methods, has shown good performance in the presence of unbalanced class distributions. In this paper, an attempt has been made to improve the prediction accuracy by using KNN weighting techniques. In this proposed method, the KNN algorithm is combined with two new weighting methods, and the results of the studies show that these methods increase the prediction accuracy compared to simple KNN and other machine learning methods. This paper presents a solution for improving preventive maintenance systems in pipelines in the oil industry.

Keywords: machine learning ،maintenance management ،oil pipeline ، weighted KNN algorithm

[1] Ph.D. Candidate, Department of Industrial Management, Science and Research Branch, Islamic Azad University, Tehran, Iran.

[2] Assistant Professor, Department of Industrial Management and Technology, Science and Research Branch, Islamic Azad University, Tehran, Iran

*Corresponding Author: E-mail: ahmad.ebrahimi@srbiau.ac.ir

[3] Associate Professor, Department of Industrial Engineering, Firouzkouh Branch, Islamic Azad University, Firouzkouh, Iran rashidi@azad.ac.ir

1. Introduction

In the midst of the Fourth Industrial Revolution, industries are constantly looking for ways to optimize production lines while also seeking to reduce their costs (Popov et al., 2021). Maintenance costs typically account for more than one-third of a manufacturing company's operating costs (Fu et al., 2020). Traditional maintenance techniques are based on two different strategies: corrective maintenance and preventive maintenance. Corrective maintenance is performed to repair faulty systems and equipment only when a failure occurs, thus reducing direct process costs (Kerf et al., 2020). But preventive maintenance is performed after regular intervals to prevent equipment and systems from failing. Therefore, repairs are performed on machines or components when they have an uncertain remaining useful life, leading to both machine downtime and increased operating costs (Zheng et al., 2021). In the oil industry, pipes play an important role in transporting oil and petroleum products, so their efficient and safe operation is crucial to minimizing environmental risk and company assets. Therefore, predicting pipeline failure is very important and necessary. Therefore, studies to predict pipeline failures have increased significantly in recent years. The present study, considering the existing shortcomings and lack of research in the field of predicting failures in oil pipelines, seeks to provide a model to identify and detect failures in oil pipelines. Machine learning techniques have been used in many studies for fault detection or failure detection and have shown good performance. Research has also been conducted on failure detection in the oil industry or failure prediction in similar equipment, but in the Iranian oil industry, research that can predict the causes of failure in oil pipes is still rarely seen. In addition, the use of a combination of metaheuristic algorithms and machine learning can achieve more accurate results in prediction by reducing the prediction error, and the use of this methodological

approach is also limited in research in the field of failure detection in oil pipes (Zenisk et al., 2019). Given the existing shortcomings and lack of research in the field of oil pipe failure prediction, this article seeks to present a model for identifying and detecting oil pipe failures, which is a serious need in the oil industry. The present study attempts to develop an appropriate model for predicting oil pipe failures with the help of machine learning algorithms and predict this failure appropriately. At the end of the article, we seek to answer this key question: How is the detection of defects in oil pipes by controlling the identified parameters with the help of machine learning algorithms?

2. Literature Review

Machine learning techniques have been used in many studies for fault detection or failure detection and have shown good performance. Research has also been conducted on failure detection in the oil industry or failure prediction in similar equipment, but in the Iranian oil industry, research that can predict the causes of failure in oil pipes is still rarely seen. In addition, the use of a combination of metaheuristic algorithms and machine learning can achieve more accurate results in prediction by reducing the prediction error, and the use of this methodological approach is also limited in research in the field of failure detection in oil pipes (Zenisk et al., 2019). Given the existing shortcomings and lack of research in the field of oil pipeline failure prediction, this article seeks to present a model for identifying and detecting oil pipeline failures, which is a serious need in the oil industry. The present study attempts to develop an appropriate model for predicting oil pipeline failures with the help of machine learning algorithms and predict this failure appropriately. At the end of the article, we seek to answer this key question: how to detect defects in oil pipelines by controlling the parameters identified with the help of machine learning algorithms?

Oil and gas pipelines are one of the most critical industrial infrastructures around the world. Failure or leakage in these lines can

lead to environmental problems, high repair costs, and even safety threats. Therefore, developing accurate predictive models that can identify failures before they occur is of great importance.

One of the challenges in predicting pipeline failures is data imbalance. In this study, various techniques were used to solve this problem. The KNN model was implemented with weighting to increase the influence of closer samples. Also, the k value was optimized so that the model could better identify minority class samples. This method improved the accuracy of the model in detecting failure samples compared to simple KNN (Atcher et al., 2020).

Traditional methods are designed based on statistical analysis and probabilistic models. Among these methods, logistic regression can be mentioned. This method has been used to predict the probability of failure, but its accuracy is low in nonlinear data. Linear discriminant analysis and quadratic discriminant analysis have also been used to classify the condition of pipes, but they have poor performance against complex data. The main drawback of traditional methods is the inability to process large volumes of industrial sensor data and the inability to model complex and nonlinear patterns, as well as sensitivity to noise and lack of training data (Megana-Mora et al., 2019).

Recent advances in machine learning have made it possible to analyze large data sets and discover complex patterns. Various studies have shown that machine learning-based methods have higher accuracy in failure prediction, some of the most important of which are:

Random Forest A model based on combining multiple decision trees that has high accuracy, but in some cases overfitting.

• Support Vector Machine (SVM) requires a lot of processing time, especially for data with certain boundaries, but in large data volumes.

XGBoost is one of the boosting methods that are capable of modeling nonlinear and complex relationships and has shown better performance than other models in many forecasting problems.

• Artificial neural networks have the ability to learn complex features from data, but they require a large amount of data for training and have low interpretability.

Therefore, the most important limitations of machine learning methods are the need for balanced data to prevent model bias

and the high computational complexity of some models. (Yang et al., 2019).

In recent years, various methods have been proposed for pipeline failure prediction. Past studies have mainly focused on the use of basic models such as Logistic Regression, LDA, and QDA. However, new research shows that ensemble learning methods such as Random Forest and Boosting can improve the prediction accuracy. This paper aims to combine the advantages of traditional algorithms and optimized KNN methods.

KNN is one of the simplest but most effective machine learning algorithms for classification and prediction problems. This method uses data spacing for decision making and has high generalizability. The most important strengths of the KNN method are its simplicity of implementation, no need for complex training, and good performance on small and unbalanced datasets. In order to improve the accuracy and performance and overcome the limitations of this method, a combination of weights has been used.

The challenges in previous methods are the lack of attention to data imbalance, which in many previous studies have run machine learning models on balanced data, while failures rarely occur in industrial data, and the lack of use of KNN optimization techniques: Previous papers have usually used standard KNN and have not investigated improvements such as weighting. (Lin et al., 2019).

In this paper, the KNN model is combined with two new weighting methods:

(a) Exponential Weighting increases the influence of closer samples and improves the accuracy of the model.

(b) Hyperbolic Tangent Weighting allows more distant samples to have an impact on the prediction, but their role is reduced.

Therefore, to further investigate the topic, the performance and accuracy of weighted KNN models are compared with machine learning methods to determine whether the proposed method can be applied in real conditions.

In order to state the problem of the paper, given that traditional methods had many limitations and could not process complex data, and since machine learning methods have higher accuracy, but require optimization for unbalanced data. Therefore, the present paper improves KNN with weighting techniques and compares it with machine learning models.

3. Methodology

The data collection tool in the present study includes a database to extract data values and a Delphi questionnaire to confirm the input variables. The Delphi questionnaire includes 8 variables mentioned in the conceptual model and the variables section. The Delphi questionnaire of the present study is set on a scale of 1 to 10 and is graded from 1 to 10 from least important to most important. The validity of the questionnaire is collected using the opinions of ten professors and its reliability is collected using the Cronbach's alpha test. The statistical population of the present study includes all experts in oil projects in the Iranian Oil Company who are preferably familiar with oil pipe failures and have technical expertise. Considering the judgmental nature of the sample, selecting 10 to 20 people leads to sample adequacy, and in the present study, at least 10 people are selected as a sample. In the above graph, the final average of each variable was obtained using the Delphi method. The most important variables are viscosity and sludge weight, followed by flow acceleration. Of course, it should be noted that the average of the variables is not very important in the present study, and only the lack of disagreement and agreement of the experts on the research variables is important, which of course was achieved. Next, we enter the machine

learning stage to predict the effect of the final eight variables using the Delphi method. Next, we identified eight variables that are effective in pipe failure. Then, based on the identified variables, it was determined to what extent these variables can be good predictors for detecting defects in oil pipes. To overcome the problem of data imbalance, oversampling techniques such as SMOTE and ADASYN have been used. These methods increase the balance in the data distribution by generating artificial samples from the minority class and improve the accuracy of the model.

In this paper, the KNN model has been optimized in two ways: exponential weighting and hyperbolic tangent. These methods have improved the prediction accuracy. To examine the impact of these techniques, the proposed models have been compared with simple KNN and other machine learning algorithms including Random Forest, SVM and XGBoost.

- Exponential Weighting in KNN
- Hyperbolic tangent weighting in KNN

In this method, the weight of each neighbor is adjusted by the hyperbolic tangent function. exponential and hyperbolic tangent graphs, which show that as the sample distance increases, the assigned weight decreases more sharply, which will have a positive effect on the performance of the method.

3. Result

In this article, in order to select the best method for predicting oil pipeline failures based on the identified variables and applying classification methods, a Python program has been used. The data is analyzed based on 8 variables and 319 samples for class analysis using different algorithms, and the results are presented.. After applying filters to pre-process the data, various machine learning methods were implemented and compared in terms of evaluation indicators in Table 2.

5. Discussion

In this paper, an optimized model for KNN was proposed, which included exponential weighting and hyperbolic tangent techniques. These techniques have increased the accuracy of the model, which indicates the practical application of this model in the oil industry. The results of this study showed that the use of weighting techniques in KNN improves the accuracy of oil pipeline failure prediction. The KNN model with exponential weighting performed better than the simple KNN, while the hyperbolic tangent weighting was superior in some indicators such as (Precision). In general, ensemble learning methods such as Random Forest and XGBoost have the best performance compared to other methods such as Logistic Regression, LDA and QDA, SVM, and KNN with hyperbolic tangent has better accuracy in Precision because its gentler weighting has increased the stability of the model. KNN with exponential weighting performs better than simple KNN because closer samples help more in decision making. Random Forest and XGBoost methods have the highest accuracy because they use a combination of multiple models to reduce variance. SVM performs better than simple KNN but is weaker than ensemble learning models because it only considers a linear decision boundary.

Considering that the research findings show that 80% of oil pipe failures can be measured using the 8 variables under study, the remaining 20% can also be investigated and identified, in other words, it can be found out what other variables can predict oil pipe failures and add them to the existing list. Of course, this work should be done according to the opinions of experts and specialists in the field under study and the final variables should be extracted.

In laboratory conditions, each of the variables can be changed and its effect on pipe failure can be further investigated. In addition, the subject of the present study was only oil pipes, and other pipes such as gas transmission pipes can also be examined and scrutinized according to the existing variables.

The rate of pipe failure is among the issues that were not included in the present study. In fact, the extent to which these variables can lead to

pipe failure is the result obtained in the present study, but the level of failure, type of failure, crack or corrosion can also be examined and considered as a variable, which is included in the limitations of the present study.

Therefore, suggestions for future research are as follows.

1- Investigate the combination of weighting methods with collective learning and meta-heuristic optimization algorithms such as genetic algorithms.

2- Consider the rate or severity of oil pipe failure.

3- Change each of the variables and examine their effect in laboratory conditions on pipe failure.

4- Examine the effect of existing variables on other types of pipes such as gas transmission pipes.

((مقاله پژوهشی))

jpm.2026.1192591/10.71737

بررسی روش‌های مبتنی بر یادگیری ماشین به منظور بهبود مدیریت نگهداری و تعمیرات در خطوط لوله نفتی با تاکید بر KNN وزن دهی شده

محمدرضا زمانی[۱]، احمد ابراهیمی*[۲]، علیرضا رشیدی کمیجان[۳]

چکیده

هزینه های تعمیر و نگهداری معمولا بیش از یک سوم هزینه های عملیاتی یک شرکت تولیدی را شامل می شود بنابراین یک استراتژی و مدل تعمیر و نگهداری بهینه می بایست باعث بهبود وضعیت عملکرد تجهیزات گردد لذا پیش بینی علل خرابی و ریشه های آن یکی از دغدغه های مهم در میان مدیران صنایع مختلف به شمار می رود. در واقع مدیران حوزه نگهداری و تعمیرات همواره به دنبال کشف این نکته هستند که چه عواملی می تواند منجر به شکست در تولید و استمرار خدمات گردد تا بدین وسیله مانع از رخداد آنها شوند .این مهم در صنایع فرایندی نظیر نفت و گاز به سبب حجم بالای ضرر و زیان ناشی از شکست در تجهیزات و توقف در تولید و انتقال محصول می تواند دارای اهمیت بیشتری نیز باشد.پیش‌بینی شکست خطوط لوله و نامتعادل بودن داده‌ها است یکی از چالش‌های اصلی در صنعت نفت است. مدل‌های سنتی قادر به شناسایی دقیق خرابی‌ها نیستند. الگوریتم کا نزدیکترین همسایگی، به عنوان یکی از روش های یادگیری ماشین در صورت وجود توزیع غیر متعادل کلاس ها عملکرد خوبی را نشان داده است . در این مقاله، با استفاده از تکنیک‌های وزن‌دهی KNN ، تلاش شده است تا دقت پیش‌بینی بهبود یابد، در این روش پیشنهادی الگوریتم KNN با دو روش نوین وزن‌دهی ترکیب شده و نتایج بررسی‌ها نشان می‌دهد که این روش‌ها دقت پیش‌بینی را نسبت به KNN ساده و سایر روشهای یادگیری ماشین افزایش می‌دهند. این مقاله راهکاری برای بهبود سیستم‌های نگهداری پیشگیرانه در خطوط لوله در صنعت نفت ارائه می‌دهد.

کلمات کلیدی: یادگیری ماشین، خطوط لوله نفتی، مدیریت نگهداری و تعمیرات، الگوریتم KNN وزن دهی شد

[۱] دانشجوی دکتری گروه مدیریت صنعتی،واحد علوم و تحقیقات دانشگاه آزاد اسلامی،تهران، ایران.
mohammadreza.zamani2229@iau.ir
[۲] استادیارگروه مدیریت صنعتی وتکنولوژی ،واحد علوم و تحقیقات دانشگاه آزاد اسلامی، تهران، ایران.(نویسنده مسئول)
ahmad.ebrahimi@srbiau.ac.ir
[۳] دانشیارگروه مهندسی صنایع،واحد فیروزکوه، ،دانشگاه آزاد اسلامی ،فیروز کوه، ایران. rashidi@azad.ac.ir

مقدمه

در میانه انقلاب صنعتی چهارم صنایع دائما به دنبال روشهایی برای بهینه سازی خطوط تولید بوده و در عین حال در صدد کاهش هزینه های خود می باشند(پوپوف[۱] ،۲۰۲۱، ۱٤). هزینه های تعمیر و نگهداری معمولا بیش از یک سوم هزینه های عملیاتی یک شرکت تولیدی را شامل می شود(فو[۲] ،۲۰۲۰، ۲). تکنیکهای سنتی تعمیر و نگهداری بر اساس دو استراتژی متفاوت است. تعمیر و نگهداری اصلاحی و تعمیر و نگهداری پیشگیرانه. تعمیر و نگهداری اصلاحی به منظور تعمیر سیستمها و تجهیزات معیوب صرفا در زمانی اجرا میشوند که یک شکست رخ داده و به این صورت هزینه های مستقیم فرایند کاهش می یابد(کرف[۳] و همکاران،۲۰۲۰، ۲۱). اما نگهداری پیشگیرانه پس از بازه های زمانی منظم جهت پیشگیری از شکست تجهیزات و سیستمها اجرا می شود. بنابراین تعمیرات در مورد ماشین آلات یا قطعات در زمانی اجرا می شود که آنها دارای عمر مفید باقیمانده نامشخص برای هم بیکاری ماشین آلات و هم افزایش در هزینه های عملیاتی می باشند(ژنگ[٤] و همکاران،۲۰۲۱، ۱۵). در صنعت نفت لوله ها نقش مهمی جهت انتقال نفت و فراورده های نفتی دارند به گونه ای که عملیات کارامد و ایمن آنها به منظور حداقل سازی ریسک محیط زیستی و دارائی های شرکت بسیار حیاتی به شمار می رود بنابراین پیش بینی شکست خطوط لوله بسیار مهم و ضروری است. متاسفانه پیش بینی شکست به دلیل مکانیسم شکل گیری یک چالش بزرگ و پیچیده در صنعت نفت است که تا کنون به طور کامل شناسایی نشده است. لذا مطالعات جهت پیش بینی شکست خطوط لوله در سال های اخیر به طور قابل توجهی افزایش یافته است. تحقیق حاضربا توجه به نقائص موجود و کمبود تحقیقات در حوزه پیش بینی شکست در لوله های نفتی به دنبال ارائه مدلی جهت شناسائی و کشف شکست در لوله های نفتی می باشد.

یک استراتژی تعمیر و نگهداری می بایست باعث بهبود وضعیت سلامت تجهیزات شده و در اداه می بایست باعث کاهش نرخ شکست تجهیزات ، حداقل سازی هزینه تعمیر و نگهداری و حداکثر سازی عمر مفید تجهیزات می باشند

---

۱ Popov

۲ Fu, H.

۳ Kerf

٤ Zheng, J

در صنعت نفت لوله های نفتی نیز به عنوان تجهیزات و ابزارهای مهمی جهت انتقال نفت به شمار می روند به گونه ای که عملیات کارامد و ایمن آنها به منظور حداقل سازی ریسک محیط زیستی انسانی و دارائی های ساختاری بسیار حیاتی به شمار می رود . اما مشکلاتی در این حوزه وجود دارد که نمی توان آنرا نادیده گرفت برای مثال شرایط متغیر عملیاتی و جریانات چند فازی در تولید نفت کنترل سیستمهای پیچیده جریان را دشوار می نماید(ابودوشیشی۱ و همکاران،۱۲،۲۰۲۱). در حال حاضر اکثریت خروجی حجیم نفتی از طریق لوله های نفتی منتقل می شود. بنابراین حفاظت از لوله های نفتی در این صنعت بسیار حائز اهمیت است چرا که نفتکشها صرفا قادر به انتقال حجم معینی از نفت بوده و سهم بیشتری از آن لوله های نفتی می باشد. به عنوان نمونه می توان گفت دریای شمال دارای ۱۵٦۷ لوله نفتی به طول ۲۵ هزار کیلومتر می باشد و البته تا سال ۲۰۰۰ تعداد ٥٤۲ رویداد شکست لوله در این دریا مشاهده شده است(الشیخ۲ و همکاران،۱۳،۲۰۱۹). تقریبا ۳۸ درصد شکست لوله های نفتی به سبب آسیب از سوی طرف سوم رخداده و ۳٦ درصد ناشی از فرسایش در خدمات بوده است. نشست لوله ها در لوله های انتقال دهنده هیدروکربن باعث مسائل مهمی از جمله خطرات زیست محیطی و زیانهای مالی هنگفت شده است.تراوش نفت خام در یک محیط اقیانوسی باعث آتش سوزی و انفجار شده است

بازرسی منظم و کنترل فرایند ابزارهای اصلی شناسائی شکست در سیستمهای خط لوله به شمار می رود محققان بسیاری از داده های گذشته برای توسعه مدلهای احتمالی در تحلیل ریسک بهره برده و تکنیکهایی نظیر مدلسازی بیزین، الگوریتمهای استنتاج فازی به کار رفته است. تحلیل پسا شکست ابزار قدرتمند دیگری برای شناسائی علل ریشه ای شکست پس از رخداد و استفاده از درسهای فراگرفته در پروژه های آتی به شمار می رود(عباس۳ و همکاران،۵،۲۰۱۹). کنترل فرایند پارامترهای جریان در طی عملیات به پرسنل به شناسائی اختلالات مهم کمک می کند اما دارای قابلیت محدودی در تمایز تفاوتهای دقیقه ای در شرایط عملیاتی می باشد. نشستی های بزرگ و کوچک اغلب تنها در زمان بازرسی روزمره صورت می گیرد که تنها زمانی موثر است که هزینه منابع و محیط عملیاتی به این مهم اجازه دهد.

---

۱ Al-Dushaishi MF,

۲ Alshaikh A

۳ K. Abbas A,

تکنیکهای یادگیری ماشین در تحقیقات بسیاری به منظور کشف خطا یا کشف شکست استفاده شده و عملکرد مناسبی را از خود نشان داده است. در خصوص کشف شکست در صنعت نفت و یا پیش بینی شکست در تجهیزات مشابه نیز تحقیقاتی صورت گرفته است اما در صنعت نفت ایران هنوز تحقیقی که بتواند علل شکست در لوله های نفتی را پیش بینی کند به ندرت مشاهده می شود. ضمن اینکه استفاده از ترکیب الگوریتمهای فراابتکاری و یادگیری ماشین می تواند با کاهش خطای پیش بینی به نتایج دقیقتری در خصوص پیش بینی دست یابد که استفاده از این رویکرد روش شناسانه نیز در تحقیقات در حوزه کشف شکست در لوله های نفتی به شکل محدودی وجود دارد(زنیسک۱ و همکاران،۲۰۱۹،۱٤).

مقاله حاضر با توجه به نقائص موجود و کمبود تحقیقات در حوزه پیش بینی شکست در لوله های نفتی به دنبال ارائه مدلی جهت شناسائی و کشف شکست در لوله های نفتی می باشد که این یک نیاز جدی در صنعت نفت به شمار می رود. تحقیق حاضر تلاش دارد با کمک الگوریتمهای یادگیری ماشین به مدلی برازنده در حوزه پیش بینی شکست در لوله های نفتی پرداخته و این شکست را به شکل مناسب پیش بینی کند. در پایان مقاله به دنبال پاسخ به این پرسش کلیدی هستیم که کشف عیوب در لوله های نفتی به وسیله کنترل پارامتر های شناسایی شده با کمک الگوریتمهای یادگیری ماشین چگونه است؟

خطوط لوله نفت و گاز یکی از حیاتی‌ترین زیرساخت‌های صنعتی در سراسر جهان هستند. شکست یا نشتی در این خطوط می‌تواند منجر به مشکلات زیست‌محیطی، هزینه‌های بالای تعمیرات و حتی تهدیدات ایمنی شود. بنابراین، توسعه مدل‌های پیش‌بینی‌کننده دقیق که بتوانند خرابی‌ها را قبل از وقوع شناسایی کنند، اهمیت زیادی دارد.یکی از چالش‌های موجود در پیش‌بینی شکست خطوط لوله، نامتعادل بودن داده‌ها است. در این پژوهش، برای حل این مشکل، از تکنیک‌های مختلفی استفاده شده است. مدل KNN با وزن‌دهی اجرا شده تا تأثیر نمونه‌های نزدیک‌تر افزایش یابد. همچنین، مقدار k بهینه‌سازی شد تا مدل بتواند نمونه‌های کلاس اقلیت را بهتر شناسایی کند. این روش باعث شد که دقت مدل در تشخیص نمونه‌های خرابی نسبت به KNN ساده بهبود یابد(اتچر۲ و همکاران،۲۰۲۰،۱۰).

Zenisek, J۱

Otchere,۲

روش‌های سنتی بر مبنای تحلیل آماری و مدل‌های احتمالاتی طراحی شده‌اند. از جمله این روش‌ها می‌توان به رگرسیون لجستیک اشاره نمود این روش برای پیش‌بینی احتمال وقوع شکست استفاده شده است، اما دقت آن در داده‌های غیرخطی پایین است . همچنین تحلیل تفکیک خطی و تحلیل تفکیک درجه دوم برای طبقه‌بندی وضعیت لوله‌ها به کار گرفته شده‌اند، اما در برابر داده‌های پیچیده عملکرد ضعیفی دارند. مشکل اصلی روش‌های سنتی عدم توانایی در پردازش حجم بالای داده‌های سنسورهای صنعتی وناتوانی در مدل‌سازی الگوهای پیچیده و غیراخطی و همچنین حساسیت به نویز و کمبود داده‌های آموزشی است(مگانا مورا[۱] وهمکاران،۲۰۱۹،۸).

پیشرفت‌های اخیر در یادگیری ماشین، امکان تحلیل داده‌های حجیم و کشف الگوهای پیچیده را فراهم کرده است. تحقیقات مختلف نشان داده‌اند که روش‌های مبتنی بر یادگیری ماشین دقت بالاتری در پیش‌بینی خرابی دارند که برخی از مهم‌ترین این روش‌ها عبارت‌اند از:

- جنگل تصادفی مدلی مبتنی بر ترکیب چندین درخت تصمیم که دقت بالایی دارد، اما در برخی موارد بیش‌برازش (Overfitting) می‌شود .
- ماشین بردار پشتیبان (SVM) به‌ویژه برای داده‌های دارای مرزهای مشخص، اما در حجم داده‌های بالا، زمان پردازش زیادی نیاز دارد.
- XGBoost از روش‌های تقویتی که قادر به مدل‌سازی روابط غیرخطی و پیچیده هستند و در بسیاری از مسائل پیش‌بینی عملکرد بهتری از سایر مدل‌ها نشان داده‌اند
- شبکه‌های عصبی مصنوعی امکان یادگیری ویژگی‌های پیچیده از داده‌ها را دارند، اما نیازمند حجم بالای داده برای آموزش هستند و تفسیرپذیری پایینی دارند.

لذا مهمترین محدودیت‌های روش‌های یادگیری ماشین نیاز به داده‌های متوازن برای جلوگیری از سوگیری مدل و پیچیدگی محاسباتی بالا در برخی از مدل‌ها است. (یانگ۲ و همکاران،۲۰۱۹،۷).

در سال‌های اخیر، روش‌های مختلفی برای پیش‌بینی خرابی خطوط لوله ارائه شده است. مطالعات گذشته عمدتاً بر استفاده از مدل‌های پایه‌ای مانند Logistic Regression، LDA و QDA متمرکز بوده‌اند. با

---

۱ Magana-Mora A,

۲ Yang,

این حال، تحقیقات جدید نشان می‌دهند که روش‌های یادگیری جمعی نظیر Random Forest و Boostingمی‌توانند دقت پیش‌بینی را بهبود بخشند. این مقاله با هدف ترکیب مزایای الگوریتم‌های سنتی و روش‌های بهینه‌سازی شده KNN انجام شده است

KNN یکی از ساده‌ترین اما مؤثرترین الگوریتم‌های یادگیری ماشین برای مسائل طبقه‌بندی و پیش‌بینی است. این روش از فاصله‌گذاری بین داده‌ها برای تصمیم‌گیری استفاده می‌کند و قابلیت تعمیم‌پذیری بالایی دارد.مهم‌ترین نقاط قوت روش KNN سادگی پیاده‌سازی و عدم نیاز به آموزش پیچیده و عملکرد مناسب در مجموعه داده‌های کوچک و نا متعادل است و به منظور بهبود دقت و عملکرد و رفع محدودیت های این روش از ترکیب وزن دهی استفاده شده است.

چالش‌های موجود در روش‌های قبلی عدم توجه به نامتعادل بودن داده‌ها می باشد که در بسیاری از مطالعات قبلی مدل‌های یادگیری ماشین را روی داده‌های متعادل اجرا کرده‌اند، درحالی‌که در داده‌های صنعتی خرابی‌ها به‌ندرت رخ می‌دهند و عدم استفاده از تکنیک‌های بهینه‌سازی :KNN مقالات قبلی معمولاً از KNN استاندارد استفاده کرده‌اند و بهبودهایی مانند وزن‌دهی را بررسی نکرده‌اند. (لین ۱ و همکاران،۲۰۱۹،۱۸).

در این مقاله، مدل KNN با دو روش نوین وزن‌دهی ترکیب شده است:

- وزن‌دهی نمایی (Exponential Weighting): تأثیر نمونه‌های نزدیک‌تر را افزایش می‌دهد و دقت مدل را بهبود می‌بخشد.
- وزن‌دهی تانژانت هایپربولیک (Hyperbolic Tangent Weighting): باعث می‌شود نمونه‌های دورتر نیز در پیش‌بینی تأثیر داشته باشند اما نقش آن‌ها کاهش یابد.

لذا جهت بررسی بیشترموضوع، عملکرد و دقت مدل‌های KNN وزن دهی شده با روش‌های یادگیری ماشین مقایسه شده‌اند تا مشخص شود آیا روش پیشنهادی می‌تواند در شرایط واقعی کاربردی باشد.

به منظور بیان مساله مقاله باتوجه به اینکه روش‌های سنتی دارای محدودیت‌های زیادی بودند و نمی‌توانند داده‌های پیچیده را پردازش کنند و از آنجا که روش‌های یادگیری ماشین دقت بالاتری دارند، اما نیازمند

---

lin,۱

بهینه‌سازی برای داده‌های نامتعادل هستند. لذا مقاله حاضر KNN را با تکنیک‌های وزن‌دهی بهبود داده و آن را با مدل‌های یادگیری ماشین مقایسه کرده است.

ابزار و روش

یکی از تجهیزات مهم که در انتقال نفت نقش حیاتی را در صنعت نفت ایفا می کند لوله های نفتی می باشد که شناسائی علل نقص در آنها با توجه به محیطها و موقعیتهای مختلف متفاوت بوده و لذا دارای اهمیت بسیاری می باشد. در خصوص شناسائی این علل در داخل کشور بر اساس داده های متقن و مسبوق به سابقه و اصطلاحات داده های تاریخی کار چندان جامعی صورت نگرفته واین در حالی است که داده های گذشته می تواند مبنای تصمیم گیری در آینده بر شمرده شود.

براین اساس تحقیق حاضر تلاش دارد مدلی جهت کشف شکست در لوله های نفتی با کمک ترکیبی از تکنیکهای یادگیری ماشین ارائه نماید. دستاوردهای تحقیق حاضر از نظر علمی می تواند به غنای ادبیات موجود در خصوص کشف عیب در تجهیزات تولیدی و به خصوص تجهیزات نفتی افزوده و بستر تحقیقات بیشتر در این حوزه را هموار نماید ضمن اینکه از نظر کاربردی می تواند ارائه الگویی در صنعت نفت باشد که مدیران این صنعت بتوانند در راستای شناسائی علل شکست لوله های نفتی از آن بهره گیری نمایند.

مقاله حاضر از نظر هدف کاربردی و از نظر جمع اوری اطلاعات توصیفی تحلیلی می باشد. مراحل انجام تحقیق به شرح ذیل است

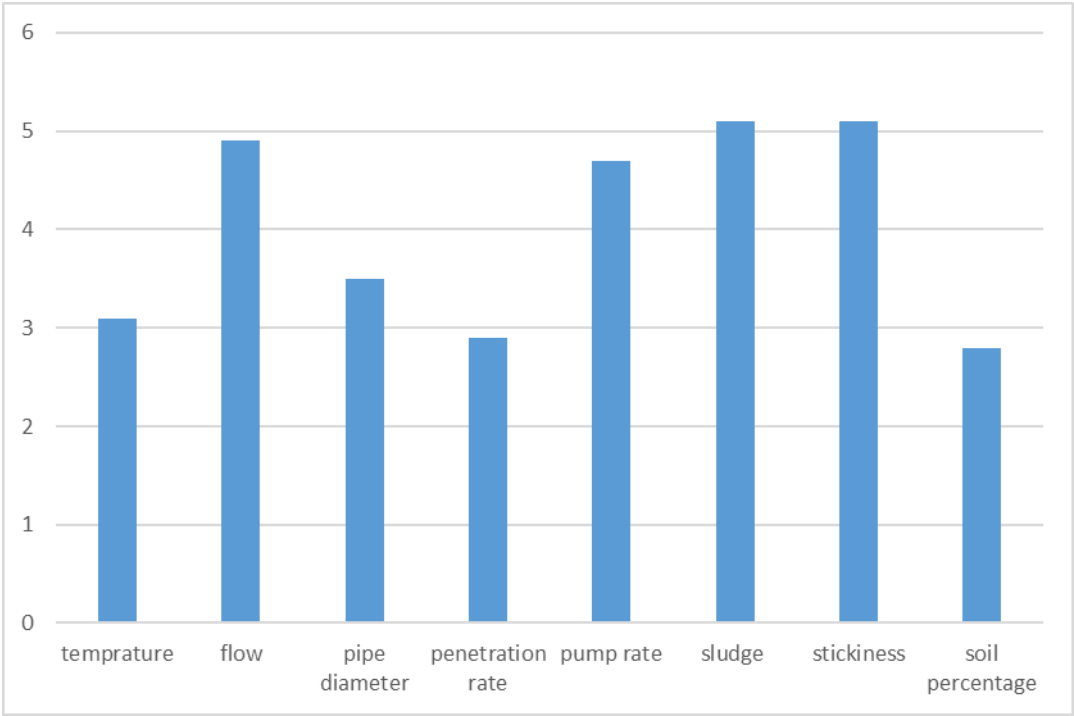
شکل شماره ۱: مدل مفهومی پژوهش

Figure 1: Conceptual model of research

در این تحقیق ابتدا مطالعات کتابخانه ای صورت گرفته و سپس بر اساس مطالعات انجام شده شکاف تحقیقاتی و نواوری استخراج می شود. سپس متغیرهای ورودی برای کشف شکست در لوله های نفتی با استفاده از رویکرد دلفی تعیین شده و در ادامه داده ها بر اساس دیتا ست تعیین شده استخراج می شود. پس از آن پیاده سازی الگوریتمهای یادگیری ماشینی جهت پیش بینی شکست در لوله های نفتی صورت می گیرد. از آن جایی که متغیرهای محدودی به شکل بسیار پیچیده در وقوع شکست تاثیرگذار بوده لذا شناسایی الگوی مبتنی بر استفاده از روش های یادگیری ماشین برای پیاده سازی در پیش بینی شکست بسیار موثر می باشد با توجه به اینکه داده های شکست دارای توزیع نامتعادل می باشد به این معنی که تعداد نمونه های شکست در مقایسه با سایر موارد بسیا کم بوده و عملکرد طبقه بندی بسیاری از روش های مبتنی بر KNN به طور چشمگیری به علت مشکل بالا، به ویژه در نمونه دارای اندازه کوچک و داده ها ی با توزیع ناهموار کاهش می یابد

ابزار جمع اوری اطلاعات در تحقیق حاضر شامل پایگاه داده برای استخراج مقادیر داده ها و پرسشنامه دلفی برای تائید متغیرهای ورودی می باشد. پرسشنامه دلفی شامل ۸ متغیر اشاره شده در مدل مفهومی و بخش متغیرها می باشد. پرسشنامه دلفی تحقیق حاضر در مقیاس ۱ تا ۱۰ تنظیم شده و از ۱ تا ۱۰ از کم اهمیت تا بسیار مهم درجه بندی شده است. روایی پرسشنامه با استفاده از نظر ده نفراز اساتید و پایایی آن با کمک آزمون آلفای کرونباخ جمع اوری می شود

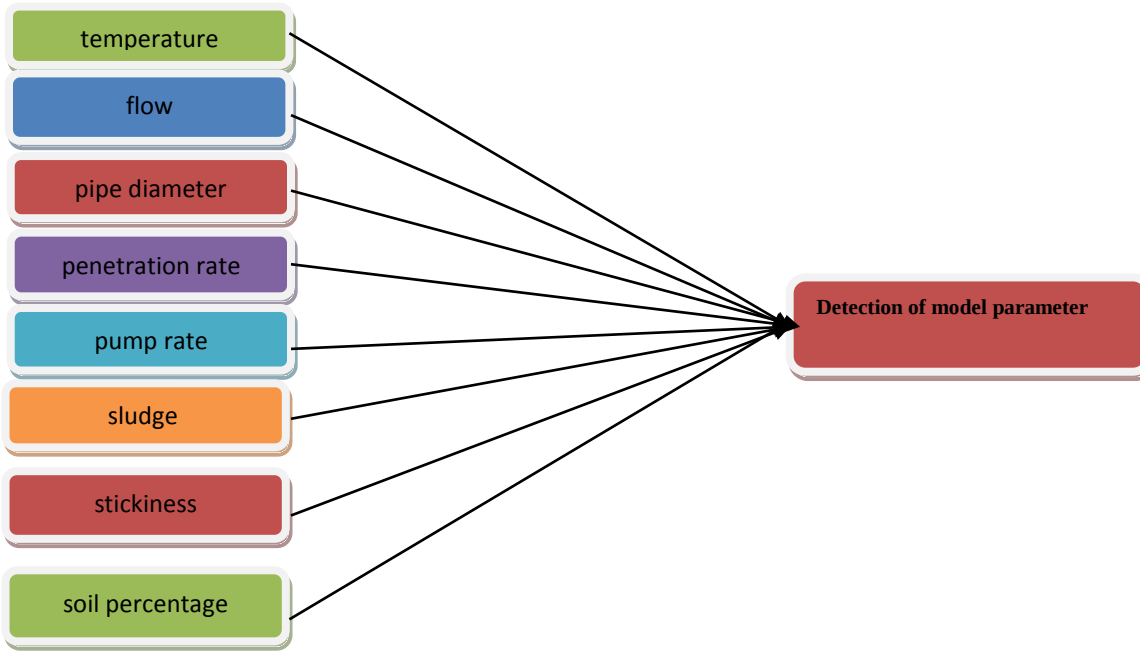
جامعه آماری تحقیق حاضر شامل کلیه خبرگان در پروژه های نفتی در شرکت نفت ایران می باشند که ترجیحا آشنا به شکست لوله های نفتی و دارای تخصص فنی باشند. با توجه به ماهیت قضاوتی نمونه انتخاب ۱۰ تا ۲۰ نفر منجر به کفایت نمونه می شود که در تحقیق حاضر حداقل ۱۰ نفر به عنوان نمونه انتخاب می شود.



شکل شماره 2: شناسایی پارامترهای با روش دلفی

Figure 2: Final parameter of Delphi method

در نمودار فوق میانگین نهایی هر یک از متغیرها با استفاده از روش دلفی حاصل شده است. مهمترین متغیرها چسبندگی و وزن لجن و پس از آن شتاب جریان می باشند. البته باید اشاره شود که میانگین متغیرها در تحقیق حاضر چندان حائز اهمیت نبوده و صرفا عدم مغایرت و توافق نظر خبرگان در خصوص متغیرهای تحقیق حائز اهمیت می باشد که البته حاصل شده است. در ادامه وارد مرحله یادگیری ماشین برای پیش بینی اثر متغیرهای هشت گانه نهایی از روش دلفی می شویم



شکل شماره ۳ شناسایی پارامترهای مدل

Figure 3: Detection of model parameter

در ادامه تعداد هشت متغیر موثر در شکست لوله را شناسایی کردیم. سپس بر اساس متغیرهای شناسائی شده تعیین شود که تا چه میزان این متغیرها می توانند پیش بینی کننده های خوبی برای کشف عیوب در لوله های نفتی باشد.

جدول شماره ۱: متغیر های مدل

Table 1: Model parameter

ردیف	Parameter title(عنوان متغیر)	Parameter type(نوع متغیر)	Parameter(متغیر)
1	Temperature(دما)	input	X1
2	flow(فشار جریان)	input	X2
3	pipe diameter(قطر لوله)	input	X3

ردیف	Parameter title(عنوان متغیر)	Parameter type(نوع متغیر)	Parameter(متغیر)
4	penetration rate(نرخ نفوذ)	input	X4
5	pump rate(نرخ پمپ)	input	X5
6	sludge (لجن)	input	X6
7	stickiness(چسبندگی)	input	X7
8	soil percentage(درصد خاک)	input	X8
9	**Pipe failure(شکست لوله)**	**output**	**Y1**

برای رفع مشکل نامتعادل بودن داده‌ها، از تکنیک‌های Oversampling نظیر SMOTE و ADASYNاستفاده شده است. این روش‌ها با تولید نمونه‌های مصنوعی از کلاس اقلیت، باعث افزایش تعادل در توزیع داده‌ها شده و دقت مدل را بهبود می‌بخشند.

در این مقاله، مدل KNN به دو روش بهینه‌سازی شده است: وزن‌دهی نمایی و تانژانت هایپربولیک. این روش‌ها باعث بهبود دقت پیش‌بینی شده‌اند. برای بررسی تأثیر این تکنیک‌ها، مدل‌های پیشنهادی با KNN ساده و سایر الگوریتم‌های یادگیری ماشین شامل Random Forest، SVM و XGBoost مقایسه شده اند.

- **وزن‌دهی نمایی در KNN**

در روش وزن‌دهی نمایی، نمونه‌های نزدیک‌تر تأثیر بیشتری در پیش‌بینی دارند. فرمول محاسبه وزن به‌صورت زیر است:

- $$w_{i=}\begin{cases}\exp\left(-\alpha\left[\frac{dk-di}{dk-d1}\times\frac{dk+d1}{dk+di}\right]^{r}\right), & dk\neq d1\\ 1, & dk=d1\end{cases}$$

dk : دورترین فاصله از نمونه مورد نظر است

که α\alphaα پارامتر تنظیمی برای تعیین میزان تأثیر فاصله است

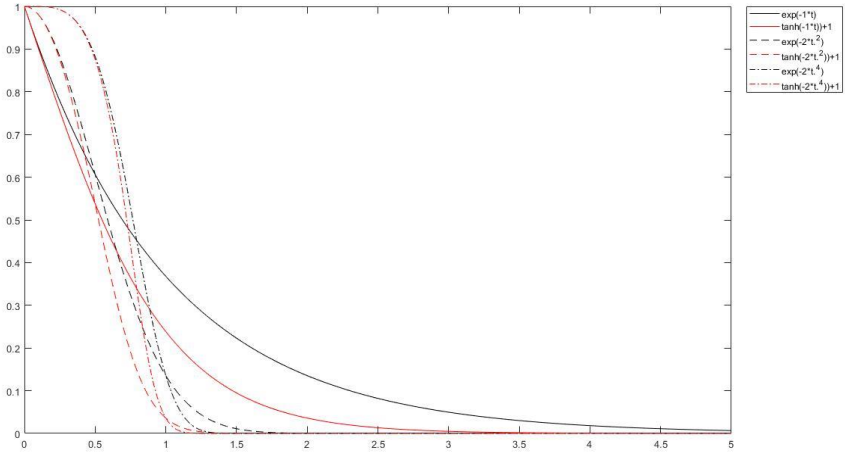
- **وزن‌دهی تانژانت هایپربولیک در KNN**

در این روش، وزن هر همسایه با تابع تانژانت هایپربولیک تنظیم می‌شود:

- $$w_{i=}\begin{cases}\tanh\left(-\beta\left[\frac{dk-di}{dk-d1}\times\frac{dk+d1}{dk+di}\right]^{r}\right)+1, & dk\neq d1\\ 1 & dk=d1\end{cases}$$

که در آن β\beta یک پارامتر کنترلی برای تعیین میزان تأثیر همسایگان است.

که در شکل زیر نمودار های نمایی و تانژانت هایپربولیک رسم شده که نشان میدهد هر چه فاصله نمو نه ها افزایش می یابد با شدت بیشتری وزن تخصیص یافته کاهش یافته و تاثیر مثبتی بر روی عملکرد روش خواهد داشت.



شکل شماره ٤: نمودار توابع نمایی و تانژانت هایپربولیک

Figure 4: Graph of exponential and hyperbolic tangent functions

- با افزایش توان r (آرگومان داخل پرانتز)، مقدار تابع (وزن) برای هر دو تابع در مقادیر پایین بیشتر و در مقادیر بالا کمتر می شود. لذا با افزایش توان آرگومان داخل تابع که در مقاله همان فرمول وزن WKNN می باشد، نتایج منطقی است که بهبود یابد.
- با افزایش ضریب آرگومان داخل پرانتز مقدار تابع (وزن) برای هر دو تابع در مقادیر پایین بیشتر و در مقادیر بالا کمتر می شود.

- با افزایش توان T، گستره تابع بیشتر شده و به تابع سیگموید شبیه تر می شود.
- با مقایسه نمودار ها مشاهده می گردد در کل، تابع tanh عملکردی بهتر از تابع نمایی exp دارد.

بهینه‌سازی هایپرپارامترها

همچنین به منظور بهینه‌سازی هایپرپارامترها، از روش‌های جستجوی شبکه‌ای و جستجوی تصادفی استفاده شده و برای الگوریتم KNN مقادیر مختلفی برای K مورد آزمایش قرار گرفت تا بهترین مقدار انتخاب شود(K=9). و روش‌های مختلف وزن‌دهی مانند وزن‌دهی یکنواخت و وزن‌دهی بر اساس فاصله بررسی شدند.

در الگوریتم SVM کرنل‌های مختلفی مانند خطی وچندجمله‌ای و تابع پایه شعاعی مورد آزمایش قرار گرفتند. وپارامتر گاما برای کرنل RBF، مقادیر مختلف گاما مورد آزمایش قرار گرفت.

در الگوریتم Random Forest مقادیر مختلفی برای تعداد درختها بررسی شدند. ومحدودیت‌های مختلفی برای عمق درخت‌ها اعمال شد تا از بیش‌برازش جلوگیری شود.

و نهایتا در الگوریتم XGBoost مقادیرمختلفی برای نرخ یادگیری و تعداد تکرارها اعمال شده است

پس از انجام بهینه‌سازی هایپرپارامترها، مدل‌ها با بهترین ترکیب هایپرپارامترها آموزش داده شدند و عملکرد آن‌ها با استفاده از معیارهای مختلف ارزیابی شد.این بهینه‌سازی منجر به بهبود قابل‌توجهی در دقت و کارایی مدل‌ها شد.

یافته ها

در این مقاله به منظور انتخاب بهترین روش جهت پیش بینی شکست خطوط لوله نفتی بر اساس متغیر های شناسایی شده و بکارگیری روش های Classification از برنامه پایتون استفاده شده است ،که داده ها بر اساس ۸ متغیر جهت بررسی کلاس به کمک الگوریتم های مختلف بررسی و نتایج ارائه می شود.

جدول شماره 2: مقایسه الگوریتم یادگیری ماشین از نظر شاخص های ارزیابی

Table 2: Comparison of machine learning algorithms in terms of performance indicators

Method روش	**Accuracy** صحت	**Precision** دقت	**Recall** بازیابی	**F1-Score** امتیاز–اف وان
KNN classic KNN کلاسیک	74.5%	68.2%	65.4%	66.7%
KNN weighted hyperbolic tangent KNN وزندهی با تانژانت هایپربولیک	**78.9%**	**74.1%**	**70.3%**	**71.2%**
KNN weighted exponential KNN وزندهی با تابع نمایی	77.3%	73.0%	68.9%	70.9%
XGBoost	**82.5%**	**81%**	**80.0%**	**79.1%**
SVM ماشین بردار پشتیبان	76.5%	74.0%	72.0%	73.0%
Random Forest جنگل تصادفی	81.2%	79.8%	78.5%	79.1%

پس اعمال فیلترها به منظور pre-processing داده ها ، روشهایمختلف یادگیری ماشین اجرا شده و از نظر شاخص های ارزیابی در جدول شماره ۲ مقایسه شده است.

جدول شماره ۳: ماتریس درهم ریختگی االگوریتم KNN

Table 3: Confusion matrix of KNN algorithm

XGBoost. ایکس جی بوست	Actual Positive مثبت واقعی	Actual Negative منفی واقعی
Predicted Positive پیش بینی مثبت	78	5
Predicted Negative پیش بینی منفی	7	110

ماتریس‌های درهم‌ریختگی هر یک از این مدل‌های KNN، SVM ، Random Forest و XGBoost بررسی شد و XGBoost با دقت بالاتر نسبت به سایر روش ها عمل کرده است. همچنین Precision روش XGBoost کمترین تعداد False Positive (FP): 0 نمونه‌ای که واقعاً منفی بوده اما مدل آن‌ها را به‌اشتباه مثبت پیش‌بینی کرده است را دارد، بنابراین دقت مثبت بالاتری دارد.

این نتایج نشان می‌دهد که روش XGBoost در این مجموعه داده‌ها عملکرد بهتری نسبت به سایر مدل‌ها دارد.

نتیجه گیری و پیشنهاد ها

در این مقاله، یک مدل بهینه شده برای KNN ارائه شد که شامل تکنیک‌های وزن‌دهی نمایی و تانژانت هایپربولیک بود. این روش‌ها دقت مدل را افزایش داده‌اند، که نشان‌دهنده کاربرد عملی این مدل در صنعت نفت است.

نتایج این مطالعه نشان داد که استفاده از تکنیک‌های وزن‌دهی در KNN باعث بهبود دقت پیش‌بینی شکست خطوط لوله نفتی می‌شود. مدل KNN با وزن‌دهی نمایی عملکرد بهتری نسبت به KNN ساده ارائه داد و SVM عملکردی بهتر از KNN ساده دارد اما نسبت به مدل‌های یادگیری جمعی ضعیف‌تر است زیرا فقط یک مرز تصمیم‌گیری خطی را در نظر می‌گیرد.درحالی‌که وزن‌دهی تانژانت هایپربولیک در تقریبا اکثر شاخص‌ها نسبت به روشهای سنتی برتری داشته است. در مجموع، روش‌های یادگیری جمعی مانند Random Forest و XGBoost بهترین عملکرد را در مقایسه با سایر روشها مانند Logistic

Regression، LDA و QDA، SVM را دارند زیرا از ترکیب چندین مدل برای کاهش واریانس استفاده می‌کنند و **KNN با تانژانت هایپربولیک Precision مناسبی را پس از روش های یادگیری جمعی داشته است چون وزن‌دهی ملایم‌تر آن باعث افزایش ثبات مدل شده است.**

با توجه به اینکه یافته های تحقیق نشان می دهد تقریبا ۸۰ درصد شکست لوله های نفتی با استفاده از ۸ متغیر مورد بررسی قابلیت سنجش است لذا ۲۰ درصد باقیمانده نیز می تواند مورد بررسی و شناخت قرار گیرد به عبارت دیگر می توان دریافت چه متغیرهای دیگری می توانند شکست لوله های نفتی را پیش بینی کرده و به فهرست موجود اضافه شوند. البته این کار می بایست با توجه به نظر کارشناسان و متخصصین امر در حوزه مورد مطالعه صورت گیرد و متغیرهای نهایی استخراج شوند.

در شرایط آزمایشگاهی می توان به تغییر هر یک از متغیرها پرداخت و اثر آنرا بر شکست لوله در ادامه بررسی نمود. ضمن اینکه مورد مطالعه تحقیق حاضر صرفا لوله های نفتی بوده است که سایر لوله ها نظیر لوله های انتقال گاز نیز می تواند با توجه به متغیرهای موجود مورد بررسی و مداقه قرار گیرد.

میزان شکست لوله در زمره مواردی بوده که در تحقیق حاضر لحاظ نشده است در واقع اینکه این متغیرها تا چه میزان می توانند منجر به شکست لوله شوند نتایجی است که در تحقیق حاضر حاصل شده است اما سطح شکست ، نوع شکست، ترک یا خوردگی نیز می تواند به عنوان یک متغیر مورد بررسی و توجه قرار گیرد که در دسته محدودیتهای تحقیق حاضر قرار دارد.

بنابراین پیشنهادات برای تحقیق آتی به شرح ذیل بیان می شود.

۱- ترکیب روش‌های وزن‌دهی با یادگیری جمعی و الگوریتم‌های بهینه‌سازی فراابتکاری مانند الگوریتم ژنتیک جهت محاسبه وزن بهینه بررسی شود
۲- در نظر گرفتن میزان یا شدت شکست لوله های نفتی
۳- تغییر هر یک از متغیرها و بررسی اثر آنها در شرایط آزمایشگاهی بر شکست لوله
٤- بررسی اثر متغیرهای موجود بر سایر انواع لوله ها نظیر لوله های انتقال گاز

تعارض منافع

نویسندگان هیچگونه تعارض منافع ندارند

References

Popov, E., Cheremisin, A., Rafieepour, S., 2021. Prediction of dead oil viscosity: machine learning vs. classical correlations. Energies 14, 930.

Lin , huang 2020 ,advanced engineering informatics ,Anovel approach for prediction forecast via improved KNN algorithm, J. Pet. Sci. Eng. 221.

Fu, H., Yang, L., Liang, H., Wang, S., Ling, K., 2020. Diagnosis of the single leakage in the fluid pipeline through experimental study and CFD simulation. J. Pet. Sci. Eng. 193

Kerf, T.D., Gladines, J., Sels, S., Vanlanduit, S., 2020. Oil spill detection using machine learning and infrared images. Remote Sens. 12, 4090.

Zheng, J., Du, J., Liang, Y., Liao, Q., Li, Z., Zhang, H., Wu, Y., 2021. Deeppipe: a semi- supervised learning for operating condition recognition of multi-product pipelines. Process Saf. Environ. Prot. 150, 510–521.

Yang, L., Fu, H., Liang, H., Wang, Y., Han, G., Ling, K., 2019. Detection of pipeline blockage using lab experiment and computational fluid dynamic simulation. J. Pet. Sci. Eng. 183.

Otchere, D.A., Ganat, T.O.A., Gholami, R., Syahrir, R., 2020. Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: comparative analysis of ANN and SVM models. J. Pet. Sci. Eng.

Al-Dushaishi MF, Abbas AK, Alsaba M, Abbas H, Dawood J (2020) Data-driven stuck pipe prediction and remedies. Upstream Oil and Gas Technol 6:1–9. https://d oi. org/ 10. 1016/j. upstre.2 020. 100024

Alshaikh A, Magana-Mora A, Gharbi SA, Al-Yami A (2019). Machine learning for detecting stuck pipe incidents: data analytics and models evaluation. Society of Petroleum Engineers, International Petroleum Technology Conference, Beijing, China. https:// doi. org/ 10. 2523/ IPTC-19394- MS

Magana-Mora A, Gharbi S, Alshaikh A, Al-Yami A (2019). AccuPipePred: a framework for the accurate and early detection of stuck pipe for real-time drilling operations. In: S.P.E. Middle East Oil and Gas Show and Conference, Manama, Bahrain. https:// doi. org/ 10. 2118/ 194980- MS

Zenisek, J.; Holzinger, F.; A_enzeller, M. Machine learning based concept drift detection for predictive maintenance. Comput. Ind. Eng. 2019, 137

Pier Francesco Orrù 1, Andrea Zoccheddu 1, Lorenzo Sassu 2, Carmine Mattia 2 , Riccardo Cozza 3 and Simone Arena(2020), Machine Learning Approach Using MLP and SVM Algorithms for the Fault Prediction of a Centrifugal Pump in the Oil and Gas Industry J. Pet. Sci. Eng. 142.

Samuel Dixon, Timothy Doster, Tegan Emerson, "To fail or not to fail: an exploration of machine learning techniques for predictive maintenance," Proc. SPIE 12113, Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications IV, 1211320 (6 June 2022); doi: 10.1117/12.2619040

Kun Yang1,2 and Xia Wang(2021), Abnormal identification of lubricating oil parameters an evaluation of physical and chemical properties based on machine learning, IOP Conference Series: Materials Science and Engineering J. Pet. Sci. Eng. 108.