

## **ОЦЕНКА СОСТОЯНИЯ СМАЗКИ ПОДШИПНИКОВ КАЧЕНИЯ С ПРИМЕНЕНИЕМ АЛГОРИТМОВ КЛАССИФИКАЦИИ**

*Целью данной работы является исследование способа решения проблемы снижения срока службы подшипников качения, установленных на промышленном оборудовании, в результате их некачественной смазки в процессе эксплуатации. Известно, что до 50% всех внеплановых простоев промышленного оборудования происходит по причине разрушения подшипников [1]. При этом основной причиной отказа подшипников являются разного рода проблемы со смазкой: избыток и недостаток смазки в подшипнике и составляют до 40% от общего числа отказов [2].*

*В работе рассматривается метод классификации состояний составе комплекса по контролю за техническим состоянием оборудования, расширяя подшипников в зависимости от их диагностических параметров: различных показателей вибрации и температуры. Классификация осуществляется методами классического машинного обучения с целью выявления наилучших показателей точности модели.*

*Разработанный алгоритм возможно успешно использовать в возможности диагностических мероприятий.*

*Классификация, машинное обучение, метод опорных векторов SVM, случайный лес RFC, k-ближайших соседей KNN, точность, подшипник, смазка.*

Krinitzin Pavel Gennadievich, Chentsov Sergey Vasilievich

## **ASSESSMENT OF BEARING LUBRICATION CONDITION WITH THE USE OF NEURAL NETWORK CLASSIFICATION**

*The purpose of this work is to explore a method for solving the problem of reduced service life of rolling bearings installed on industrial equipment due to their poor lubrication during operation. It is known that up to 50% of all unplanned downtime of industrial equipment occurs due to bearing failure [1]. At the same time, the main cause of bearing failure is various lubrication problems: excess and lack of grease in the bearing account for up to 36% of all failures [2].*

*This work examines a method for classifying the conditions of bearings depending on their diagnostic parameters: various indicators of vibration and temperature. Classification is carried out using traditional machine learning methods with the aim of identifying the best model accuracy indicators.*

*The developed algorithm can be successfully used as part of a system for monitoring the technical condition of equipment, thereby expanding the capabilities of diagnostic measures.*

*Classification, machine learning, Support Vector Machine (SVM) method, Random Forest Classifier (RFC), k-nearest neighbors (KNN), accuracy, bearing, lubrication.*

### **Введение**

По данным современных исследований, до 40% прибыли теряет в среднем промышленное предприятие, по причине внеплановых простоев оборудования [1]. Учитывая, что состояние смазки подшипников играет в этих простоях доминантную роль, необходимо обеспечить непрерывный контроль за состоянием смазки подшипников в процессе эксплуатации. Смазка в подшипнике выполняет следующие функции [4]:

1. Образует между рабочими поверхностями упругую гидродинамическую пленку, защищающую тела качения, кольца и сепаратор от повреждений;

2. Снижает трение скольжения между телами качения подшипника;

3. Защищает подшипник от попадания загрязнений и образования коррозии;

4. Способствует эффективному охлаждению подшипника, способствуя тепло распределению.

Количество смазки в подшипниках, к сожалению, зачастую зависит от человеческого фактора и того, как сервисный персонал произведет техническое обслуживание оборудования.

Второй проблемой, требующей сокращения внеплановых отказов подшипников, является текущий дефицит качественных изделий на рынке. Это связано с уходом с рынка крупнейших производителей подшипниковой продукции – SKF, FAG, Timken, NSK, что связано с санкционной политикой иностранных производителей. По состоянию на 2009г. соотношение производителей подшипников на мировом рынке представлено на рис.1 [5]:



*Рис. 1. Распределение производителей подшипников на мировом рынке*

Предлагаемая модель позволяет выполнять диагностическую функцию по контролю за состоянием смазки в подшипниках в период эксплуатации, тем самым увеличивая ресурс их работы.

#### **Основная часть**

Для проведения исследования использовался набор данных «Vibration analysis metrics of a ball bearing during different operational states», находящийся в свободном доступе на сайте Google datasetsearch. Маркированный набор данных предоставляет собой файл формата CSV и содержит показатели работы подшипника качения 6204. К ним отно-

сят характеристики вибрации: виброскорость ( $v$ -RMS), виброускорение ( $a$ -RMS), ударное виброускорение ( $a$ -Peak).

Целевая переменная – bearing state, характеризует работу подшипника в следующих режимах:

- 1) шарикоподшипник с уплотнением с рекомендуемым количеством смазки;
- 2) шарикоподшипник без уплотнения без смазки;
- 3) шарикоподшипник без уплотнения с избыточным количеством смазки.

В процессе исследования датасета был выполнен статистический анализ признаков для каждого значения целевого класса. Выявлены и удалены из обучающей выборки признаки с значительными отклонениями от среднестатистических показателей – выбросы (рис.2).

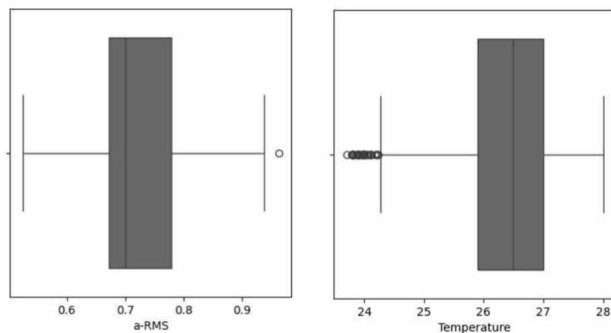


Рис. 2. Диаграммы рассеяния *Boxplot*

Был выявлен разброс в значениях признаков, характеризующих температуру и вибрационные показатели, отличающиеся в 5 раз между собой, что потребовало произвести нормализацию данных с помощью библиотеки *StandardScaler*.

Отнесение целевой переменной к одному из заранее определенных классов — это задача классификации, поэтому для ее решения были использованы следующие базовые алгоритмы [6]:

1. Метод *KNN*;
2. *RandomForestClassifier*;
3. Метод *SVM* - метод опорных векторов.

Метод *K*-ближайших соседей (*K-Nearest Neighbors*) основан на поиске кратчайшей дистанции между тестируемым объектом и ближай-

шими к нему классифицированным объектами из обучающего набора данных [7] (1):

$$a(u) = \operatorname{argmax}_{y \in Y} \sum_{i=1}^m [y(x_{i,u}) = y] * w(i, u), \quad (1)$$

где  $u$  – объект классификации,  $y$  – целевой класс,  $w(i, u)$  – весовая функция оценивающая степень важности  $i$ -го соседа.

$$w(i, u) = [i \leq k], \quad (2)$$

где  $k$  – целочисленное значение, характеризующее количество соседей (2).

В методе KNN выбор значения  $k$  очень важен. При маленьком значении  $k$  "шум" будет оказывать сильное воздействие на итоговый результат. Большое значение делает его затратным с точки зрения вычислений и противоречит основной идее KNN, что близкие точки могут иметь схожие плотности или классы [8]. Значение по умолчанию для количества  $k$ -соседей - 5 ближайших соседей; Экспериментальным путем определено значение  $K = 4$ , при котором достигается минимальный процент ошибок модели.

В качестве метрики модели использовано расстояние Минковского [9] (3):

$$p(x, y) = (\sum_{i=2}^n |x_i - y_i|^p)^{\frac{1}{p}} \quad (3)$$

где  $x, y$  – точки пространства, параметр  $p$  принят равным 2.

Классифицируемый объект будет относиться к тому классу, к которому принадлежит ближайший объект набора. Определение наиболее влиятельных признаков на работу модели производилось с помощью дисперсионного анализа по методу Anova [10]. Точность классификации по методу KNN составила 98,6% и 98,1% на тренировочных и тестовых значениях соответственно.

RandomForestClassifier (RFC) - этот классификатор разбивает данные на всё меньшие и меньшие подмножества на основе разных критериев, т. е. у каждого подмножества своя сортирующая категория. С каждым разделением количество объектов определённого критерия уменьшается [11]. Классификация подойдёт к концу, когда сеть дойдёт до подмножества только с одним объектом (4):

$$a(x) = \frac{1}{N} \sum_{i=1}^N b_i(X) \quad (4)$$

где  $a(x)$  – классификатор,  $N$  – количество деревьев,  $i$  – счетчик деревьев,  $b$  – решающее дерево,  $x$  – выборка данных.

Критичным параметром для деревьев решений, влияющим на точность модели, является глубина обучения – `max_depth`. В рамках исследова-

дования параметр `max_depth` принят равным 15. Определение параметра производилось с помощью инструмента `GridSearch()` библиотеки `scikit-learn`. Анализ значимости признаков производился с помощью встроенного метода `feature_importances`.

Точность классификации по методу RFC составила 100% и 98,2% на тренировочных и тестовых значениях соответственно

SVM – принцип работы алгоритма построен на предположении, что чем больше расстояние (зазор) между разделяющей гиперплоскостью и объектами разделяемых классов, тем меньше будет средняя ошибка классификатора (5):

$$a(\vec{x}) = \text{sign}(\sum_{i=1}^c w_i X_i - b) \quad (5)$$

где  $\vec{x}$  – вектор значений признаков объекта,  $w_i, b$  – параметры гиперплоскости,  $c$  – константа.

С целью инициализации значимых признаков использован алгоритм Recursive feature elimination (RFE) – рекурсивный поиск и удаление признаков по значимости [11].

Точность классификации по методу SVM составила 97,1% и 97,3% на тренировочных и тестовых значениях соответственно

### **Результаты**

Все три модели показали удовлетворительный результат обучения и могут быть применены в качестве практической модели классификации подшипников по состоянию их смазки:

1. Модель `RandomForestClassifier` показала наиболее высокие результаты классификации как на тестовых, так и на тренировочных данных. На тестовых данных ее ошибка не превышает 2%;

2. Наиболее значимым признаком для предсказания целевой переменной "Bearing State" у всех трех моделей оказался признак `V-rms` - показатель виброскорости (м/с).

### **Выводы**

В рамках настоящей работы по разработке алгоритма диагностирования подшипников качения на предмет достаточного количества смазочных материалов в них, произведено исследование применения трех моделей классификации (метод ближайших соседей, случайный лес и метод опорных векторов). Выполнен анализ влияния признаков при работе каждой модели на предсказание целевой переменной.

В настоящее время, в промышленности все большее значение уделяется вопросам автономной диагностики технического состояния обо-

рудования. На основе результатов показаний датчиков вибрации и температуры, установленных на оборудовании, прогнозируются отказы оборудования и сроки ремонтов. Предлагаемая модель позволяет расширить возможности использования диагностических данных для решения задач оптимизации сервисного обслуживания оборудования. Что может быть обеспечено путем повышения качества выполнения операций по смазке оборудования, производя ее в достаточном количестве и в требуемое время.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Shen Zhang, Shibo Zhang, Bingnan Wang, Thomas G. Habetler*. Deep Learning Algorithms for Bearing Fault Diagnostics // IEEE Access. 2022. Vol. 8. P. 29857 – 29881.

2. *Исмаилов Р. Н., Гареев Р.Р., Ямалиев В. У.* Прогнозирование остаточного ресурса подшипника по уровню вибрации механизма // Экспозиция Нефть Газ. 2015. N3. С. 65 – 68.

3. *Антонов В. С., Везин В. А.* Борьба с простоями как способ повышения экономической эффективности бизнеса [Электронный ресурс] // Экономика и жизнь: электрон. журн. 2021. №33. URL: <https://www.eg-online.ru/article/440363> (дата обращения: 16.04.2024).

4. *Щец С. П., Сакало В. И.* Влияние смазочного материала на процессы, протекающие в подшипниках качения // Транспортное машиностроение. 2016. N2. С. 31 – 35.

5. *Дрыночкин А. В., Аленина Е. Э., Тришкин А. Г.* Анализ состояния российских предприятий подшипниковой подотрасли и смежных отраслей промышленности // Известия МГТУ «МАМИ». 2012. N2(14). С. 8 – 13.

6. *Бабаев А. М., Шемякина М. А., Ляшов М. В.* Обзор классических методов машинного обучения в контексте решения задач классификации // Форум молодых ученых. 2018. N11(27). С. 137 -142.

7. *Ульянов Н. В., Ахмедова Ш. А.* Введение в анализ данных // Актуальные проблемы авиации и космонавтики. 2022. N2. С. 357 – 359.

8. *Yeray Mezquita, Ricardo S. Alonso, Roberto Casado-Vara, Javier Prieto & Juan Manuel Corchado*. A Review of k-NN Algorithm Based on Classical and Quantum Machine Learning // Distributed Computing and Artificial Intelligence, Special Sessions, 17th International Conference. 2020. P. 189 – 198.

9. Стрюков Р.К., Шашкин А.И. О модернизации метода ближайших соседей // Вестник ВГУ, серия: системный анализ и информационные технологии. 2015. №1. С. 114 – 120.

10. Аннаева М., Атанасосова А. Однофакторный дисперсионный анализ: методы и применение в статистике // Всемирный ученый. 2023. №9(1). С. 267 – 271.

11. Борисов П. Д., Косолатов Ю.В. Способ оценки схожести программ методами машинного обучения // Труды Инст. сист. Прогр. РАН. 2022. №5(34). С. 63 – 75.

**Криницин Павел Геннадьевич**, аспирант, менеджер по надежности ООО «ИСО» филиал в г. Братск, 665729, Россия, город Братск, улица Советская 6, телефон +79501228427, email: alfa\_reklama@mail.ru.

**Ченцов Сергей Васильевич**, доктор технических наук, профессор кафедры систем автоматики, автоматизированного управления и проектирования Сибирского федерального университета, 660074, Россия, город Красноярск, улица Киренского 26 корпус 1, телефон +7 (391) 291-22-35, email: schentsov@sfu-kras.ru.

**Krinitzin Pavel Gennadievich**, postgraduate student, reliability manager at LLC "ISO" branch in Bratsk, Russia, 665729, Bratsk city, Soviet Street 6, phone +79501228427, email: alfa\_reklama@mail.ru.

**Chentsov Sergey Vasilievich**, Doctor of Technical Sciences, professor at the Department of Systems Automation, Automated Control and Design of Siberian Federal University, 660074, Russia, Krasnoyarsk city, Kirenskogo street 26, building 1, phone +7 (391) 291-22-35, email: schentsov@sfu-kras.ru.