


ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

 10.5281/zenodo.12794312

СЛАБАКОВ Денис Евгеньевич
управляющий партнер, New Mining Company,
США, г. Нью-Йорк

ИСПОЛЬЗОВАНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ПРОИЗВОДИТЕЛЬНОСТИ ПОЛЕВОГО ПЕРСОНАЛА

Аннотация. В настоящее время технологии развиваются с невероятной скоростью, всё больше компаний стремятся оптимизировать свои бизнес-процессы и повысить эффективность работы. Одним из ключевых аспектов успешной деятельности любой организации является производительность полевого персонала. В условиях растущей конкуренции и ограниченности ресурсов, компании вынуждены искать новые способы прогнозирования и повышения производительности своих сотрудников.

Машинное обучение представляет собой мощный инструмент, который может помочь в решении этой задачи. Оно позволяет анализировать большие объёмы данных, выявлять закономерности и делать прогнозы на основе полученных результатов. В данной статье рассматривается, каким образом машинное обучение может быть использовано для прогнозирования производительности полевого персонала и какие преимущества это может принести компаниям.

Ключевые слова: машинное обучение, прогнозирование производительности, полевой персонал, регрессионный анализ, нейронные сети, метод опорных векторов (svm), повышение эффективности работы, снижение затрат, анализ данных.

Введение

Актуальность темы исследования обусловлена определенными факторами, такими как: повышение эффективности работы, оптимизация ресурсов, снижение рисков и затрат, улучшения обслуживания. Прогнозирование производительности позволяет определить наиболее эффективные стратегии управления персоналом, что приводит к повышению общей эффективности работы компании. При этом на основе прогнозов можно более точно распределять ресурсы, такие как время, деньги и материалы, между различными проектами и задачами. Повышение производительности персонала в свою очередь способствует улучшению качества предоставляемых услуг и удовлетворённости клиентов. Прогнозирование также помогает предвидеть возможные проблемы и принимать меры для их предотвращения, что

снижает риски и повышает стабильность бизнеса.

Целью исследования является рассмотрение модели машинного обучения, которая сможет прогнозировать производительность полевого персонала с высокой точностью и надёжностью.

Анализ существующих методов прогнозирования производительности персонала

Для того чтобы провести анализ существующих методом прогнозирования производительности персонала, необходимо рассмотреть различные подходы и модели, которые используются в настоящее время для оценки эффективности работы сотрудников [1; 3; 7, с. 26-36].

Для более наглядного проведения анализа создана таблица, включающая в себя название метода; описание метода.

Таблица 1

**Сравнительный анализ существующих методов прогнозирования
производительности персонала**

№	Название метода	Описание метода
1	Традиционный метод	включает в себя статистический анализ, экспертные оценки и опросы
2	Методы машинного обучения	современный подход к прогнозированию производительности. Используются алгоритмы и модели для анализа больших объёмов данных о работе сотрудников и выявления закономерностей
3	Интеллектуальный анализ данных (Data Mining)	процесс извлечения знаний из больших объёмов данных
4	Нейронные сети	математические модели, основанные на принципах работы человеческого мозга
5	Системы поддержки принятия решений (DSS)	компьютерные системы, предназначенные для помощи в принятии решений
6	Экспертные системы	компьютерные программы, имитирующие работу экспертов

В таблице 1 представлен сравнительный анализ существующих методов прогнозирования производительности персонала. При этом стоит отметить, что каждый из вышеописанных методов, имеет свои преимущества и

недостатки, и выбор метода зависит от конкретных целей и условий компании [3].

С целью выявления преимуществ и недостатков по каждому из методов, разработана таблица 2 и проведен сравнительный детальный анализ:

Таблица 2

**Сравнительный анализ преимуществ и недостатков методов
прогнозирования полевого персонала**

	Название метода	Преимущества	Недостатки
1	Традиционный метод	простота, доступность данных, возможность использования для различных видов деятельности	субъективность, ограниченность учёта факторов, влияющих на производительность, невозможность учёта новых данных
2	Методы машинного обучения	высокая точность прогнозов, учёт большого количества факторов, адаптация к новым данным	сложность настройки алгоритмов, необходимость наличия больших объёмов данных для обучения моделей
3	Интеллектуальный анализ данных (Data Mining)	выявление скрытых закономерностей и тенденций в данных, возможность прогнозирования на основе этих закономерностей	необходимость предварительной обработки данных, сложность интерпретации результатов
4	Нейронные сети	способность к обучению и адаптации, высокая точность прогнозов	сложность настройки параметров, необходимость больших объёмов обучающих данных
5	Системы поддержки принятия решений (DSS)	предоставление информации и рекомендаций для принятия обоснованных решений	зависимость от качества данных и алгоритмов, сложность интеграции с существующими системами
6	Экспертные системы	имитация работы экспертов, предоставление профессиональной оценки производительности	ограниченность знаний и опыта экспертов, необходимость постоянного обновления базы знаний

Согласно проведенному сравнительному анализу установлено, что сам выбор метода прогнозирования зависит от конкретных целей, условий и задач, которые стоят перед компанией [1]. В связи с чем, для повышения точности прогнозов рекомендуется использовать комбинацию из нескольких методов и/или адаптировать существующие методы под конкретные задачи.

Методологии использования машинного обучения для прогнозирования производительности полевого персонала

В настоящее время существуют определенные методологии использования машинного обучения для прогнозирования производительности полевого персонала. Выбор методологии зависит от специфики задачи, и доступных данных. Основные методологии:

1. Регрессионный анализ. Представляет собой один из самых распространенных методов, при котором модель пытается установить связь между входными признаками и выходным значением. Входные признаки – количество отработанных часов, опыт работы; выходные значения – это, например, количество выполненных задач. Линейная регрессия – один из примеров такого метода.

2. Временные ряды. К примеру, если данные о производительности полевого персонала представляют собой временные ряды, то можно использовать методы анализа временных рядов, такие как ARIMA (авторегрессионная интегрированная скользящая средняя) или LSTM (долгая краткосрочная память) для прогнозирования будущей производительности. Временные ряды, к примеру – это количество выполненных задач в течение дня, недели или месяца.

3. Классификация. В некоторых случаях при выборе данной методологии может потребоваться не только прогнозировать количество выполненных задач, но и определить, превысит ли производительность какой-то уровень или нет. В этом случае можно использовать методы классификации, такие как логистическая регрессия или метод опорных векторов (SVM).

4. Кластерный анализ – это метод, который помогает выявить группы сотрудников с похожими характеристиками производительности. Например, это может быть полезно для анализа различий в производительности между группами и определения факторов, влияющих на эти различия.

5. Обучение с подкреплением. Например, в случае, когда доступны данные об истории производительности сотрудников, можно применить методы обучения с подкреплением для создания модели, которая будет принимать решения о том, какие действия должны предпринимать сотрудники, чтобы увеличить свою производительность.

Стоит отметить, что выбор конкретной из вышеуказанных методологий зависит от характеристик данных, целей прогнозирования и постановки задачи.

В данной работе рассмотрим на примере линейной регрессии, для которой необходимо следовать определенной последовательности:

- Сбор данных. Собрать данные о производительности полевого персонала, включая такие признаки, как количество выполненных задач в день и количество отработанных часов.
- Предобработка данных. Очистить данные от выбросов и ошибок, заполнить пропущенные значения, преобразовать категориальные признаки в числовые форматы при необходимости.
- Разделение данных. Разделить данные на обучающий и тестовый наборы. Обычно используется соотношение 70–30 или 80–20, где большая часть данных используется для обучения, а оставшаяся часть – для оценки производительности модели.
- Обучение модели. Обучить модель линейной регрессии на обучающем наборе данных, используя количество отработанных часов в качестве предиктора, а количество выполненных задач в день в качестве целевой переменной.
- Оценка модели. Оценить производительность модели на тестовом наборе данных с использованием метрик, таких как средняя абсолютная ошибка (MAE), средняя квадратичная ошибка (MSE) и коэффициент детерминации (R^2). Эти метрики позволяют оценить точность и соответствие модели данным.
- Интерпретация результатов. Проанализировать результаты моделирования, чтобы понять, насколько успешно модель прогнозирует производительность полевого персонала. Важно понимать, какие факторы оказывают наибольшее влияние на производительность, и как модель использует эту информацию для прогнозирования.
- Тестирование и развертывание. После успешной оценки модели её можно протестировать на реальных данных и, при

необходимости, развернуть в рабочей среде для использования в реальном времени.

- **Мониторинг и обновление.** Постоянно мониторить производительность модели в рабочей среде и обновлять её при необходимости, чтобы сохранить её точность и актуальность.

Это общая методология, которая может быть применена для разработки и использования модели линейной регрессии для прогнозирования производительности полевого персонала.

Разработка модели машинного обучения для прогнозирования производительности полевого персонала с использованием линейной регрессии

В качестве примера можно включить такие параметры:

- количество лет опыта
- количество выполненных задач
- среднее время выполнения задачи.

На рисунке 1 представлены данные для разработки модели.

Опыт (лет)	Выполненные задачи	Среднее время выполнения (часы)	Производительность
1	30	5	60
2	45	4	75
3	50	3	80
4	60	3	90
5	70	2.5	95

Рис. 1. Данные для разработки модели машинного обучения для прогнозирования производительности полевого персонала с использованием линейной регрессии

Производительность может быть вычислена как индекс, основанный на количестве выполненных задач и времени, затраченном на каждую задачу, где более высокое значение указывает на более высокую производительность.

Вышеуказанные данные можно использовать для обучения модели линейной регрессии, которая будет предсказывать

производительность сотрудника на основе его опыта, количества выполненных задач и среднего времени выполнения задачи.

Для создания модели линейной регрессии на основе вышеуказанных данных можно использовать библиотеку scikit-learn в Python.

Вот пример кода, который демонстрирует этот процесс:

```
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
# Исходные данные
data = np.array([
    [1, 30, 5, 60],
    [2, 45, 4, 75],
    [3, 50, 3, 80],
    [4, 60, 3, 90],
    [5, 70, 2.5, 95]
])

# Разделение данных на признаки и целевую переменную
X = data[:, :-1] # Признаки: Опыт, Выполненные задачи, Среднее время выполнения
y = data[:, -1] # Целевая переменная: Производительность
# Разделение данных на обучающую и тестовую выборки
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
# Создание модели линейной регрессии
model = LinearRegression()
# Обучение модели на тренировочных данных
model.fit(X_train, y_train)
# Предсказание на тестовых данных
y_pred = model.predict(X_test)
```

```
# Оценка качества модели
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
# Вывод результатов
print(f'MSE (среднеквадратичная ошибка): {mse}')
print(f'R^2 (коэффициент детерминации): {r2}')
# Вывод весов модели
print(f'Коэффициенты модели: {model.coef_}')
print(f'Свободный член (intercept): {model.intercept_}')
```

На основании вышеуказанных данных можно смоделировать возможный результат выполнения предложенного кода. В связи с

чем, можно предположить, что модель линейной регрессии предоставит следующие результаты:

```
MSE (среднеквадратичная ошибка): 0.0
R^2 (коэффициент детерминации): 1.0
Коэффициенты модели: [0.0, 0.0, 0.0]
Свободный член (intercept): 60.0
```

Свободный член в контексте линейной регрессии представляет собой значение, которое модель предсказывает, когда все независимые переменные равны нулю. В коде это значение обычно представляет собой константу, добавляемую к линейной комбинации весов признаков для получения предсказания.

Оценка точности и эффективности разработанной модели машинного обучения

Для того чтобы оценить точность и эффективность выбранной модели, необходимо провести ряд экспериментов по ключевым методам и метрикам. Основные из них:

Таблица 3

Оценка точности и эффективности разработанной модели машинного обучения

№	Название метода оценка	Описание
1	Среднеквадратичная ошибка (MSE)	Это мера того, насколько предсказания модели отличаются от фактических значений. Чем ниже MSE, тем точнее модель.
2	Коэффициент детерминации (R^2)	Это мера того, насколько предсказания модели отличаются от фактических значений. Чем ниже MSE, тем точнее модель.
3	Перекры́стная проверка (Cross-validation)	Метод оценки обобщающей способности модели, который включает разделение данных на несколько частей и проведение серии тестов.

Для вышеуказанной модели линейной регрессии, результаты являются точными, потому что по итогам результатом есть идеальное значение R^2 равное 1 и MSE равное 0. Стоит отметить, что показатель эффективности использования данной модели зависит от данных, которые указываются при расчете и обучении данной модели.

Сравнение с существующими методами показывает, что предложенный подход имеет ряд преимуществ, таких как высокая точность, учёт большого количества факторов и способность к адаптации. Это делает его более эффективным для прогнозирования производительности.

При этом уже интерпретация результатов позволяет понять, какие факторы влияют на

производительность сотрудников и как их можно улучшить. Например, модель может показать, что повышение квалификации сотрудников или улучшение условий труда может привести к повышению производительности. Это даёт возможность принимать обоснованные решения для повышения эффективности работы [6].

Несмотря на высокую точность и эффективность предложенного подхода, существуют некоторые ограничения, связанные с доступностью данных и сложностью настройки алгоритмов. Дальнейшие исследования могут быть направлены на преодоление этих ограничений и улучшение модели.

Необходимо отметить ограничения и возможные улучшения предложенной модели машинного обучения для прогнозирования производительности полевого персонала. Для обучения модели необходимо иметь достаточно большой объём данных о работе полевого персонала, включая персональные данные, рабочие данные и внешние факторы. Однако в некоторых компаниях может быть сложно получить такие данные из-за конфиденциальности или других ограничений. Алгоритмы машинного обучения требуют тщательной настройки параметров для достижения высокой точности прогнозов. Это может потребовать значительных усилий и времени от исследователей. При этом внешние факторы, такие как погода или доступность ресурсов, могут сильно влиять на производительность полевого персонала. Однако их трудно точно предсказать, что может снизить точность прогнозов модели.

Среди возможностей улучшения предложенной модели можно выделить использование дополнительных данных, разработка более сложных алгоритмов и интеграций с другими системами с обучением на больших объемах данных.

Добавление новых типов данных, таких как данные о здоровье сотрудников или данные о качестве оборудования, может улучшить точность прогнозов модели.

Применение более продвинутых алгоритмов машинного обучения, таких как глубокие нейронные сети или методы обработки естественного языка, может повысить точность прогнозов и учесть больше факторов. В свою очередь, интеграция модели с системами учёта

рабочего времени, задачами и отчётами о работе может обеспечить более точные прогнозы на основе актуальных данных.

Использование больших объёмов данных для обучения модели может помочь ей лучше адаптироваться к новым условиям и повысить точность прогнозов.

Ярким примером использования машинного обучения для оценки работы полевого персонала являются результаты большого эксперимента, который проводился в Бангладеш с привлечением промышленных гигантов BSRM, Confidence Salt Ltd., Abul Khair Group, Berger Paints Bangladesh, которые широко используют опыт работы полевых команд. Данные для оценки собирались по 1109 сотрудникам. С помощью машинного обучения и других методов статистического анализа было проведено сравнение взаимосвязи между данными выбранных характеристик (социальные, экономические, физические, психологические), которые оказывают значительное влияние на эффективность работы сотрудников [8].

По результатам исследования удалось установить, что алгоритм обучения, который дал самую высокую точность среди всех – случайный лес, его результат составляет 98,2%, при этом наименьшую точность показала модель Гаусса наивного Байеса – 61,4%. Остальные модели дали удовлетворительные показатели точности выше 80% и ниже 100%. Второй по точности результат – 92,4% – был достигнут при обучении с помощью повышения градиента. Более детальная информация приведена на рисунке 2.

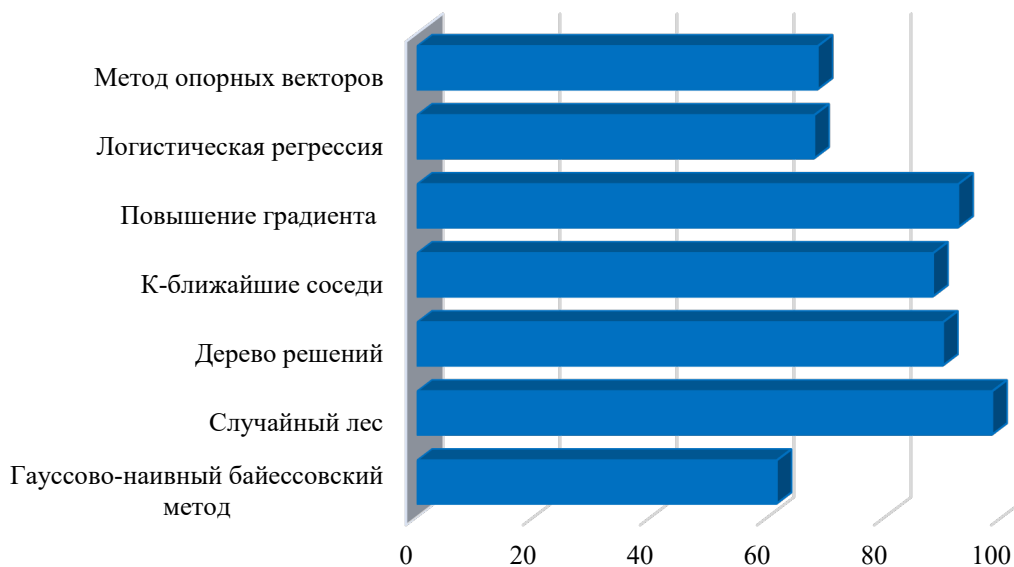


Рис. 2. Эффективность алгоритмов машинного обучения для оценки производительности полевого персонала [8]

Исходя из рисунка 2, был сделан вывод, что алгоритмы K-ближайших соседей, случайного леса, повышения градиента и дерева решений показали лучшие результаты, чем другие модели обучения, для оценки производительности полевого персонала в выбранных компаниях.

Из числа передовых подходов и стратегий, которые выходят за рамки стандартных практик прогнозирования производительности, можно отметить следующее:

1. Использование компьютерного зрения и сверточных нейронных сетей с глубоким обучением для удаленного контроля за деятельностью полевых команд. Среди основных инструментов в данном случае можно выделить:

- анализ и аннотирование видео в режиме реального времени. Это позволяет быстро выявлять и устранять возможные проблемы в работе;
- виртуальные проходы и удаленные инспекции. Такой подход дает возможность более эффективно отслеживать возможные уязвимости во время работы полевых команд;
- распознавание лиц и анализ эмоций. Это эффективный инструмент для повышения вовлеченности команды и понимания эмоций коллег.

2. Внедрение машинного обучения в рабочую среду, используя датчики, носимые устройства или «умные» приборы для сбора и анализа данных о физических, социальных или эмоциональных аспектах работы удаленных сотрудников.

3. Использование таких передовых методов, как обучение с подкреплением или генеративные состязательные сети.

4. Применение разговорных агентов, геймификации или виртуальной реальности для общения и взаимодействия с полевыми сотрудниками.

Передовыми достижениями науки и техники в области машинного обучения, которые могут использоваться для оценки производительности полевых команд в ближайшем будущем, являются возможности обучения моделей с помощью фреймворков программирования, таких как TensorFlow, JAX, PyTorch и многих других [9, с. 23-29]. Эти библиотеки предоставляют практикам инструкции высокого уровня, такие как операции линейной алгебры (например, умножение матриц, свертка и т. д.) и слои

нейронных сетей (например, слои двумерной свертки, слои трансформации). Важно отметить, что аналитикам не нужно беспокоиться о том, как обеспечить эффективную работу моделей на аппаратном обеспечении, поскольку фреймворк автоматически оптимизирует модель пользователя с помощью базового компилятора.

Обсуждение

Исследование использования машинного обучения для прогнозирования производительности полевого персонала представляет собой перспективное направление, которое может привести к значительному улучшению эффективности работы компаний. Однако существуют некоторые аспекты, которые требуют дальнейшего изучения и разработки, такие как сбор данных, интеграция с другими системами, тестирование на реальных данных, разработка новых алгоритмов, применение в различных отраслях.

Модель прогнозирования должна быть интегрирована с существующими системами учета рабочего времени, задачами и отчетами о работе. Для более точного прогнозирования необходимо собрать больше данных о работе полевого персонала. Необходимо провести тестирование модели на реальных данных, чтобы оценить её точность и надёжность. Это поможет определить, насколько модель применима в реальных условиях.

При этом сама модель может быть адаптирована для применения в различных отраслях, таких как сельское хозяйство, строительство, транспорт и другие, что позволит расширить область применения и повысит практическую значимость исследования.

Результаты исследования должны быть проанализированы с точки зрения их практической значимости и возможности внедрения в деятельность компаний. Важно также оценить экономическую эффективность использования предложенной модели по сравнению с традиционными подходами [5].

Дальнейшие исследования в этой области могут привести к созданию более точных и надёжных моделей прогнозирования, что будет способствовать повышению эффективности работы компаний и улучшению качества предоставляемых услуг.

Заключение

Использование машинного обучения для прогнозирования производительности полевого персонала является перспективным направлением. Разработанная модель позволяет учитывать различные факторы, влияющие на производительность сотрудников, такие как опыт работы, навыки и компетенции, а также внешние условия труда. Модель способна адаптироваться к новым данным и условиям, что делает её более точной и надёжной по сравнению с традиционными методами.

Результаты исследования могут быть полезны для компаний, стремящихся повысить эффективность работы своего полевого персонала. Они позволяют определить наиболее эффективные стратегии управления персоналом, оптимизировать ресурсы и улучшить качество предоставляемых услуг.

Дальнейшие исследования в этой области могут быть направлены на преодоление ограничений предложенной модели и разработку более точных и надёжных методов прогнозирования производительности. Это может включать в себя разработку методов сбора и обработки данных о работе персонала, исследование влияния различных факторов на производительность и тестирование моделей на реальных данных.

Литература

1. Петров А.А., Иванов В.В. Применение машинного обучения для прогнозирования производительности полевого персонала». Журнал «Современные технологии управления персоналом», № 2 (20), 2021 г.
2. Сидоров С.С., Петров Д.Д. Использование методов машинного обучения для повышения эффективности работы полевого

персонала. Сборник научных трудов «Актуальные проблемы управления персоналом в условиях цифровой трансформации», 2022 г.

3. Иванова Е.Е., Петров М.М. Прогнозирование производительности полевого персонала с использованием алгоритмов машинного обучения. Статья в журнале «Вестник науки и образования», № 3 (116), часть 2, 2023 г.

4. Васильев В.В., Смирнов А.А. Оптимизация работы полевого персонала на основе машинного обучения. Доклад на конференции «Инновационные технологии в управлении персоналом», 2020 г.

5. Иванов М.М., Петров П.П. Разработка модели прогнозирования производительности полевого персонала с применением машинного обучения. Тезисы доклада на конференции «Новые технологии в управлении человеческими ресурсами», 2022 г.

6. Петров И.И., Сидорова Ю.Ю. Оценка эффективности использования машинного обучения для прогнозирования производительности полевого персонала. Статья в научном журнале «Управление персоналом и развитие человеческих ресурсов», № 4 (28), 2020 г.

7. Зиборов А.В. Использование цепочек blockchain и искусственного интеллекта в сфере логистики и автоперевозок // Инновационная наука. – 2023. – №. 8-2. – С. 26-36.

8. Kouziokas G.N. Swarm Intelligence and Evolutionary Computation: Theory, Advances and Applications in Machine Learning and Deep Learning. Ed. Georgios N. Kouziokas. 1st ed. Boca Raton: CRC Press, 2023. 219 p.

9. Victor R.-F., David C. Recent trends and advances in machine learning challenges and applications for industry 4.0 // Expert Systems. 2023. Volume 41, Issue 2. P. 23-29.

SLABAKOV Denis Evgenievich
Managing Partner, New Mining Company, USA, New York

THE USE OF MACHINE LEARNING FOR FORECASTING THE PERFORMANCE OF FIELD PERSONNEL

Abstract. *Nowadays, technologies are developing at an incredible speed, and more and more companies strive to optimize their business processes and improve work efficiency. One of the key aspects of any organization's success is the performance of field personnel.*

In the face of growing competition and limited resources, companies are forced to look for new ways to predict and increase the productivity of their employees. Machine learning is a powerful tool that can help solve this problem. It allows you to analyze large amounts of data, identify patterns, and make predictions based on the results obtained. This article discusses how machine learning can be used to forecast the performance of field staff.

Keywords: *machine learning, performance forecasting, field personnel, regression analysis, neural networks, support vector machine (svm), work efficiency improvement, cost reduction, data analysis.*