

КРИВЕНЧУК ЮРІЙ

Національний університет "Львівська політехніка"

<https://orcid.org/0000-0002-2504-5833>e-mail: Yurii.P.Kryvenchuk@lpnu.ua

ВОВЧАК ЛЮБОМИР

Національний університет "Львівська політехніка"

e-mail: liubomyr.vovchak.knm.2019@lpnu.ua

ПРОГНОЗУВАННЯ ТЕНДЕНЦІЙ АВТОМОБІЛЬНОГО РИНКУ ЗА ДОПОМОГОЮ МЕТОДІВ МАШИННОГО НАВЧАННЯ

В роботі наведено результати дослідження теми прогнозування тенденцій світового автомобільного ринку та розроблення системи для розв'язання такої задачі, з метою впровадження таких систем у відповідні сфери в Україні для подальшої завчасної підготовки до потенційних змін. Виділено та описано такі основні етапи: визначення тестових наборів даних, навчання нейронної мережі, передбачення тенденцій. Після проведення аналізу результатів було виявлено, що створення системи передбачення тенденцій світового автомобільного ринку є актуальним та доцільним завданням на сьогодні, а найефективнішим інструментом для цього є використання нейронних мереж.

Ключові слова: передбачення ринку, нейронні мережі, авто, ринок, тенденції.

KRYVENCHUK YURII, VOVCHAK LIUBOMYR

Lviv Polytechnic National University

CREATION OF AUTOMOTIVE MARKET PREDICTION SYSTEM

Automotive market prediction systems are becoming increasingly important in today's realities. They are used to predict the vector of changes in the world car market for example, the increase in the share of electric cars or the growing popularity of station wagons. These systems can help manufacturers better understand the needs of the buyer and make timely changes in their strategy as needed. Automotive market forecasting systems can also help to match specific changes to the markets of different countries according to their needs. Providing a more objective approach to determining the company's business plan, which will be based on the data that these systems will provide. And of course, salary forecasting systems can be useful for companies to make a business plan and calculate their finances, having a better understanding of further market changes, the company can make an appropriate budget, taking into account the costs of developing those changes that will be popular in the future. The practical value of this work is a system of forecasting automotive market trends, which can provide valuable information and help make more informed decisions that can bring the most benefit to the company. In addition, such a system will be very useful for investors. They will be able to know in advance the trends of the world market and invest in companies that are developing in a priority direction, which will greatly reduce investment risks. And, of course, such a program will greatly facilitate the life of the buyers themselves. Now the market will be regulated much faster and as soon as there is any need for changes - companies will be ready because they have been warned about this need for changes by our forecasting system. In this way, customers will be able to receive the most suitable product on time and without unnecessary costs for emergency development.

Keywords: automotive market trends prediction, neural networks.

Постановка проблеми

Автомобільний ринок розвивається і адаптується під потреби споживача дуже швидко та динамічно і тільки системи прогнозування можуть надати цінну інформацію, яка допоможе автомобільним компаніям завчасно приймати більш обґрунтовані рішення щодо подальшого розвитку та розробки технологічних інновацій.

Автомобільні концерни зможуть використовувати систему прогнозування тенденцій, щоб обрати для себе стратегію, яка з великою імовірністю стане популярною та принесе великі прибутки. Саме така «гра на випередження» зможе дати дрібним компаніям шанс на швидкий розвиток та отримання більших прибутків.

Аналіз останніх джерел

У статті [1] дослідження зосереджено на побудові математичної моделі, яка могла б передбачити ціну вживаного автомобіля на основі його поточних характеристик. Визначення ціни вживаного автомобіля є складним завданням, оскільки кілька факторів, таких як поточний пробіг, поточний стан, марка, рік тощо, можуть впливати на прогнозовану ціну автомобіля. І з точки зору людини, яка продає, точно передбачити ціну вживаного автомобіля стає дилемою. Таким чином, суть цього виклику полягає у вивченні моделей, які можуть правильно розраховувати ціну вживаного автомобіля насамперед на основі його можливостей. Завдяки цьому, у свою чергу, споживач може зробити більш усвідомлену покупку. Тому у статті впроваджено та досліджено різні методи машинного навчання з аналізом даних, щоб надати точне та просте у використанні рішення.

У дослідженні [2] представлена модель оцінки безпеки транспортного засобу, що проїжджає по мосту великого прогону на основі повнозв'язаної нейронної мережі (FCN). Першим кроком є дослідження реакцій великого прольоту моста на вітрове збудження за допомогою випробування в аеродинамічній трубі та моделі кінцевих елементів. Згодом наводяться типові моделі транспортних засобів і встановлюється система транспортний засіб – міст, враховуючи погодні умови. Оцінено типи аварійності транспортних

засобів з несприятливими погодними умовами. Зокрема, визначаються вхідні та вихідні змінні моделі оцінки безпеки транспортного засобу, і одночасно досягаються дані навчання, валідації та тестування. Було порівняно та проаналізовано 29 моделей за допомогою прихованого шару, початкової швидкості навчання, розміру партії, функції активації та методу оптимізації. Виявлено, що модель 4-15-15-4 має найкращу продуктивність прогнозування, і вона може забезпечити своєрідну корисність для управління дорожнім рухом і зменшити ймовірність аварій транспортних засобів на мосту.

Для того, щоб ефективно сприяти сталому розвитку транспорту, необхідно визнавати реальне споживання палива транспортними засобами. У дослідженні [3] проведено порівняльний аналіз шляхом побудови п'яти регресійних моделей для прогнозування реального рівня споживання палива легкими бензиновими транспортними засобами в Китаї на основі великих даних з точки зору факторів транспортних засобів, факторів навколишнього середовища, а також факторів водіння. Результати показують, що регресійна модель випадкового лісу працює найкраще серед п'яти моделей-кандидатів із середньою абсолютною похибкою 0,630 л/100 км, середньою абсолютною відсотковою похибкою 7,5%, середньоквадратичною похибкою 0,805, квадратом R 0,776, і 10-кратний бал перехресної перевірки 0,791. Крім того, охоплено найважливіші характеристики, що впливають на споживання палива серед 25 факторів з трьох вищевказаних точок зору. Відповідно до важливості кожного фактора в найоптимальнішій моделі три найважливіші фактори — це послідовність гальмування та прискорення, потужність двигуна та свідомість власників транспортних засобів щодо економії палива.

В дослідженні [3] було доведено, що поведінка водія є одним із головних факторів, які впливають на споживання палива автомобіля. Метою роботи [5] є вдосконалення баз даних моніторингу споживання палива на основі даних мобільних телефонів. На основі терміналів мобільного телефону та бортової діагностичної системи (OBD), встановлених у таксі, витягуються дані про поведінку водіння та дані про витрату палива відповідно. Зіставляючи дані про поведінку водіння, зібрані мобільним телефоном, із даними про витрату палива, зібраними OBD, досліджується кореляція між поведінкою водіння та споживанням палива, щоб можна було передбачити споживання палива автомобіля на основі даних мобільного телефону. Моделі прогнозування споживання палива будуються з використанням нейронної мережі зворотного поширення (BP), опорної векторної регресії (SVR) і випадкових лісів. Усі три моделі могли точно прогнозувати споживання палива з абсолютною відносною похибкою менше 10%. Доведено, що модель випадкового лісу має найвищу точність і працює швидше, що робить її придатною для широкого застосування. Цей метод закладає основу для моніторингу вдосконалення бази даних і точного управління споживанням палива міським транспортом.

Оскільки саме ціна є ключовим елементом автомобільного ринку – розвинемо тему, розкрити в статті [1]. Стандартні алгоритми машинного навчання використовуються для прогнозування витрат у кількох статтях. Однак люди рідко замислюються про ефективність окремих моделей і замість цього ігнорують менш відомі, але більш складні. В статті [5] представлено багатоперіодну теоретичну модель динамічного ціноутворення. У цьому аналізі розглянуто кілька найсучасніших моделей, використовуючи як традиційні, так і найсучасніші методи машинного навчання, щоб оцінити вплив функцій на стратегії прогнозування. Як додатковий бонус, це дослідження забезпечує надійну перевірку численних стратегій впровадження регресійної моделі, що обіцяє перспективне прогнозування цін на автомобілі в майбутньому.

Мета дослідження [6] полягає у дослідницькому аналізі даних конкретного набору даних, який складається зі специфікацій автомобілів протягом діапазону років, щоб знайти взаємозв'язки між цими характеристиками та кореляціями з характеристикою ціни, а також створити модель машинного навчання прогнозування ціни з порівнянням реальних цін і прогнозованих цін наприкінці дослідження. Дослідження проводиться за допомогою набору даних, зібраного з веб-сайту турецького продавця автомобілів arabam.com за допомогою веб-скрипта, написаного на Python. Техніка лінійної регресії використовується на етапі оцінки моделі. Виявлено, що прогнозовані ціни збігаються з реальними цінами з майже 75% успіхом. Деякі обмеження застосовуються як до набору даних, так і до функції ціни, щоб отримати кращі результати для показників середньої абсолютної помилки (MAE) і середньоквадратичної помилки (RMSE).

У статті [7] наведено аналіз ДТП. У ньому представлено еволюцію систем безпеки, починаючи з опису систем, що використовуються в транспортному засобі, з особливим акцентом на передбаченні засинання водія. У статті також пропонується власна система прогнозування сну на основі розпізнавання осіб водіїв. Виявлення орієнтирів обличчя представлено як двоетапний процес: алгоритм знаходить обличчя в цілому, а потім потрібно локалізувати ключові структури обличчя в області, що цікавить. У статті представлено роботу алгоритму виявлення засинання водія; метод виявлення та аналізу.

Правильна техніка прогнозування може дозволити компаніям отримати більший дохід і спростити планування політики у виробництві. У дослідженні [8] запропоновано інтеграцію експоненціального згладжування Холта-Вінтерса з оптимізацією параметрів золотого перерізу для прогнозування продажів автомобілів в Індонезії. Метод золотого перетину використовується для знаходження оптимальних параметрів як вхідних даних для методу експоненційного згладжування Холта-Вінтерса. Отримані параметри будуть використані для прогнозування продажів автомобілів в Індонезії. На основі експериментів на кожній марці автомобіля було отримано три різні параметри для отримання оптимального значення MAPE.

Дослідження [9] в основному представляє наступні аспекти: поточна ситуація інтелектуального виробництва в автомобільному виробництві, система інтелектуального виробництва та склад інтелектуальної виробничої системи для виробництва автомобілів, структурний дизайн мережі автомобільного інтелектуального виробництва, дослідження системи позиціонування автомобільного виробництва і точки проектування системи позиціонування UWB, а також метод впровадження аналізу великих даних автомобільного інтелектуального виробництва.

Стаття [10] будує модель даних та алгоритм прийняття рішень про витрати для спеціалізованих автомобілебудівних підприємств на основі калькуляції витрат об'єкта – стандарту – діяльності (OS-ABC) метод, який забезпечує підтримку даних для нього. На основі теорії дерева рішень у відношенні до спеціалізованих автомобілебудівних підприємств, стаття будує модель дерева рішень, висуває алгоритм генерації, який може перетворити цю модель на бінарне дерево, нарешті розробляє модуль визначення вартості.

Метою роботи є створення сучасної та якісної системи передбачення тенденцій автомобільного ринку за допомогою методів машинного навчання.

Виклад основного матеріалу

Огляд наборів даних

Проаналізувавши роботи у численних джерелах, було прийнято рішення для навчання моделей нейронної мережі обрати кілька наборів даних та об'єднати їх частини. У таблиці 1 наведено основні характеристики проаналізованих датасетів.

Таблиця 1 [1]

Огляд наборів даних, що містять інформацію про продажі авто в різних країнах

База даних	Кількість прикладів	Локація	Параметри
New Car Sales in Norway	4377	Norway	Year, Month, Make, Quantity, Ptc
Car Sale Advertisements	9576	Ukraine	Car, price, body, mileage, engV, engType, registration, year, model, drive
Used Cars Dataset	427 тис.	USA	id, url, atregion, region_url, price, year, manufacturer, model, condition, cylinders
Used Car Auction Prices	559 тис	USA	year, make, model, rim, body, transmission, vin, state, condition, odometer

Етапи процесу передбачення заробітної плати

Процес прогнозування тенденцій автомобільного ринку складається з наступних етапів:

- **Збір даних.** Початковий етап складається з пошуку даних з різних джерел та вибір ключових характеристик, від яких залежать тенденції ринку.
- **Підготовка даних.** На цьому етапі ми очищуємо та попередньо обробляємо дані, щоб забезпечити їх точність і якість
- **Розробка функцій.** Під час виконання даного етапу створюються нові функції з існуючих даних.
- **Вибір моделі.** Обираємо найбільш відповідний алгоритм машинного навчання для виконання поставленого завдання, а саме нейронну мережу. Пріоритетними нейронними мережами для прогнозування є RNN, FNN та MLP. Проаналізувавши численну кількість робіт, було прийнято рішення обрати кілька архітектур рекурентних та нейронних мереж прямого зв'язку та натренувати їх, після чого на основі швидкодії та точності результатів вибрати найоптимальнішу.
- **Навчання моделі.** Використовуючи підготовлені навчальні дані, налаштовуючи гіперпараметри та оцінюючи продуктивність за допомогою таких показників, як середня квадратична помилка або середня абсолютна помилка, навчаємо обрану вище модель.
- **Оцінка моделі.** На фінальному етапі оцінюємо продуктивність розгорнутої моделі на нових даних, для визначення її точності і внесення необхідних коректив.

Висновок

В результаті проведеної роботи було створено систему прогнозування тенденцій автомобільного ринку. Ця розробка дозволяє автомобільним компаніям та інвесторам завчасно скорегувати свою стратегію та бізнес-план з урахуванням імовірних змін попиту. Детальне дослідження цієї галузі показало, що автомобільний ринок дуже швидко змінюється, а автоконцерни не мають можливості динамічно реагувати на такі зміни, тому прогнозування тенденцій автомобільного ринку є доцільним та актуальним.

Впродовж дослідження доведено, що саме методи машинного навчання є ідеальним рішенням для задач такого плану, вони показують хороші результати в цій галузі, пропонуючи високий рівень точності та здатність обробляти складні та нелінійні зв'язки в даних.

Література

1. Santosh Kumar Satapathy; Rutvikraj Vala; Shiv Virpariya. An Automated Car Price Prediction System Using Effective Machine Learning Techniques. Date of Conference: 20-21 May 2022, doi: 10.1109/CISES54857.2022.9844350
2. Yang, Lin Yang, Bo Wu, Gang Yao, Hang Li, Soltys Robert, Safety Prediction Using Vehicle Safety Evaluation Model Passing on Long-Span Bridge with Fully Connected Neural Network. Date of Conference: 15 Oct 2019, doi: <https://doi.org/10.1155/2019/8130240>
3. Yushan Yang, Nuoya Gong, Keying Xie, Qingfei Liu Predicting Gasoline Vehicle Fuel Consumption in Energy and Environmental Impact Based on Machine Learning and Multidimensional Big Data. Date of Conference: 22 February 2022, doi: <https://doi.org/10.3390/en15051602>
4. Ying Yao, Xiaohua Zhao, Chang Liu, Jian Rong, Yunlong Zhang, Zhenning Dong, Yuelong Su Vehicle Fuel Consumption Prediction Method Based on Driving Behavior Data Collected from Smartphones. Date of Conference: 23 Mar 2020, doi: <https://doi.org/10.1155/2020/9263605>
5. Fadhil Muhammad Basysyar; Ferisanti; Maryam Wulandari; Indah Sucitra; Dian Ade Kurnia; Solikin Prediction of Automobiles Prices Using Exploratory Data Analysis Based on Improved Machine Learning Techniques. Date of Conference: 08-09 December 2022, doi: <https://doi.org/10.1109/ICIC56845.2022.10006925>
6. Mert Can Satioglu; Yilmaz Ar; Bulent Tugrul Automobile Price Prediction in Turkey Marketplace with Linear Regression. Date of Conference: 21-23 October 2021, doi: <https://doi.org/10.1109/ISMSIT52890.2021.9604688>
7. Radosław Wróbel, Gustaw Sierzputowski, Piotr Haller, Veselin Mihaylov, Radostin Dimitrov Automobile Price Prediction in Turkey Marketplace with Linear Regression. Date of Conference: 17 November 2022, doi: <https://doi.org/10.19206/CE-142950>
8. Mamluatul Hani'Ah; Ika Kusumaning Putri; Ariadi Retno Tri Hayati Ririd Parameter Optimization Of Holt – Winters Exponential Smoothing Using Golden Section Method for Predicting Indonesian Car Sales. Date of Conference: 14-15 September 2021, doi: <https://doi.org/10.1109/IEIT53149.2021.9587379>
9. Hao Xu; Jinzhou Lin; Lin Liu; Tao Zhao; Jian Wu; Xinzhu Liu Research on Intelligent Manufacturing of Automobile Production Based on Big Data. Date of Conference: 18-20 November 2022, doi: <https://doi.org/10.1109/WCMEIM56910.2022.10021395>
10. Hong-tao Yang Research on Cost Decision of Specialized-Automobile Manufacturing Enterprise Based on the Theory of Decision Tree. Date of Conference: 18-20 December 2020, doi: <https://doi.org/10.1109/ICDMA.2010.464>