Methodological Article

Isfahan Artificial Intelligence Event 2023: Drug Demand Forecasting

Abstract

Background: The pharmaceutical industry has seen increased drug production by different manufacturers. Failure to recognize future needs has caused improper production and distribution of drugs throughout the supply chain of this industry. Forecasting demand is one of the basic requirements to overcome these challenges. Forecasting the demand helps the drug to be well estimated and produced at a certain time. **Methods:** Artificial intelligence (AI) technologies are suitable methods for forecasting demand. The more accurate this forecast is the better it will be to decide on the management of drug production and distribution. Isfahan AI competitions-2023 have organized a challenge to provide models for accurately predicting drug demand. In this article, we introduce this challenge and describe the proposed approaches that led to the most successful results. **Results:** A dataset of drug sales was collected in 12 pharmacies of Hamadan University of Medical Sciences. This dataset contains 8 features, including sales amount and date of purchase. Competitors compete based on this dataset to accurately forecast the volume of demand. The purpose of this challenge is to provide a model with a minimum error rate while addressing some qualitative scientific metrics. **Conclusions:** In this competition, methods based on AI were investigated. The results showed that machine learning methods are particularly useful in drug demand forecasting. Furthermore, changing the dimensions of the data features by adding the geographic features helps increase the accuracy of models.

Keywords: *Drug demand forecasting, Isfahan artificial intelligence competitions, Supply chain management*

Submitted: 04-Sep-2024 **Revised:** 18-Sep-2024 **Accepted:** 18-Sep-2024 **Published:** 23-Jan-2025

Meysam Jahani[1], Zahra Zojaji[1], AhmadReza Montazerolghaem[2], Maziar Palhang[3], Reza Ramezani[1], Ahmadreza Golkarnoor[1], Alireza Akhavan Safaei[1], Hossein Bahak[1], Pegah Saboori[1], Behnam Soufi Halaj[1], Ahmad R. Naghsh-Nilchi[4], Fatemeh Mohamadpoor[5], Saeid Jafarizadeh[4]

[1]Department of Software Engineering, Faculty of Computer Engineering, University of Isfahan, Isfahan, Iran, [2]Department of Information Technology, Faculty of Computer Engineering, University of Isfahan, Isfahan, Iran, [3]Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran, [4]Artificial Intelligence Department, Faculty of Computer Engineering, University of Isfahan, Isfahan, Iran, [5]Artificial Intelligence Department, Faculty of Computer Engineering, Shiraz University, Shiraz, Iran

Address for correspondence:
Dr. Zahra Zojaji, Department of Software Engineering, Faculty of Computer Engineering, University of Isfahan, Isfahan, Iran.
E-mail: z.zojaji@eng.ui.ac.ir

Introduction

In recent years, the pharmaceutical landscape has experienced increased drug production by various manufacturers. These drugs can save numerous lives within the healthcare field. For instance, in 2019, the availability of vaccines produced by multiple manufacturers saved many lives from COVID-19.[1] Effective management of these drugs can lead to production at scale, priced affordably, and reduce wastage unnecessarily.[2]

To enhance drug demand management efficiency, it is crucial to forecast the future needs of a drug supply chain. Forecasting is the process of estimating or projecting future demands over a certain period utilizing historical data and information, occurrences, or events. In the drug industry, demand forecasting helps to optimize processes and decision-making in different parts of the drug supply chains such as logistics, marketing, and sales.[3,4] For example, effective forecasting can help to properly inventory, reduce extra costs for an organization, and improve customer service. Forecasting future requirements enables timely and adequate production of pharmaceuticals, thereby addressing issues related to overproduction, such as drug expiration and unnecessary expenses for procurement and storage. Furthermore, forecasting demand can help prevent critical shortages of drugs.[5] Demand forecasting timelines may vary depending on the position within the supply chain and the necessity for strategic decision-making. For example, pharmacies may require daily sales forecasts to plan employee schedules, whereas manufacturers often forecast production on an annual basis.[6]

Recent technologies, such as artificial intelligence (AI), are playing a crucial role in drug demand prediction, today. These advanced techniques can be

How to cite this article: Jahani M, Zojaji Z, Montazerolghaem A, Palhang M, Ramezani R, Golkarnoor A, *et al.* Isfahan artificial intelligence event 2023: Drug demand forecasting. J Med Signals Sens 2025;15:2.

Access this article online

Website: www.jmssjournal.net

DOI: 10.4103/jmss.jmss_59_24

Quick Response Code:

broadly classified into two categories: classic methods and deep neural network-based methods. Classic methods are considered simple and conventional approaches for forecasting drug demand.[7] These methods involve building a linear model using independent variables such as date, season, hospital characteristics, and other relevant factors to predict drug consumption over the next 10 months. Techniques such as linear regression, support vector machines, decision trees, fuzzy models, and basic neural networks are commonly utilized in these classic methods. On the other hand, deep neural network-based methods are recognized as powerful tools for prediction tasks. In the realm of deep learning, recurrent neural networks (RNNs) such as long short-term memory (LSTM) and gated recurrent unit, as well as convolutional neural networks, are extensively employed. These sophisticated networks can extract intricate features from data and effectively forecast drug usage and various other prediction-related issues. However, it is essential to note that deep neural network methods necessitate substantial amounts of data and high-processing systems to operate efficiently.

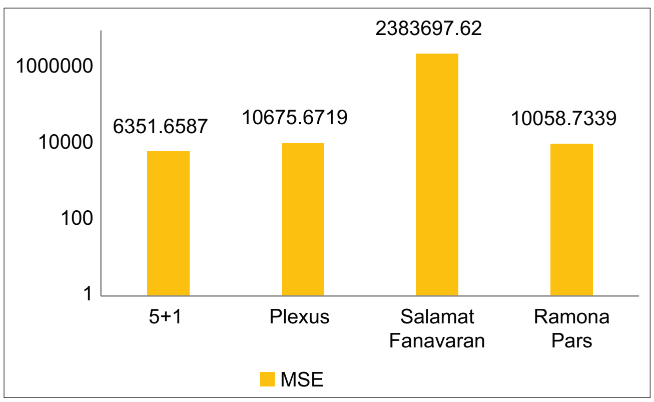

Figure 1: Compare models of demand forecasting in the first stage evaluation

Recent research has employed classical and neural network methods across various sectors of the supply chain. Some studies have proposed integrated statistical and deep learning models for drug forecasting within the phase distribution of the supply chain.[8] In addition, other methodologies have recommended inventory management based on deep reinforcement learning approaches.[9] Furthermore, in certain studies, LightGBM models have been utilized in the machine learning module to suggest top-rated or optimal medications to customers in the pharmaceutical industry.[10] These research efforts have also introduced methods to propose sales and marketing strategies by considering the trend and seasonal effects of different pharmaceutical product categories with unique characteristics.[11] However, several challenges are raised with these methods. Variations in laws and the drug supply chain across different countries, as well as differences in disease prevalence and drug consumption levels, require ongoing investigation to identify an efficient machine learning algorithm for precise drug demand forecasting to support stakeholders in supply and demand analysis. Moreover, it is imperative to assess forecasting models from a domain-specific perspective within each country, in addition to a broader view.

In this context, the Isfahan National Elite Foundation hosted a drug demand forecasting competition during the Isfahan AI (IAI)-2023 events to advance retinal forecasting approaches in drug scope through various AI-based techniques. During this competition, participants were given a dataset containing drug sales information. Researchers utilized the training data to refine their methods and subsequently submitted their algorithm outputs for evaluation using a separately provided test dataset. To ensure an unbiased assessment of performance, the test data labels were concealed, preventing any partiality related to customized method selection or parameter choices tailored to the dataset.

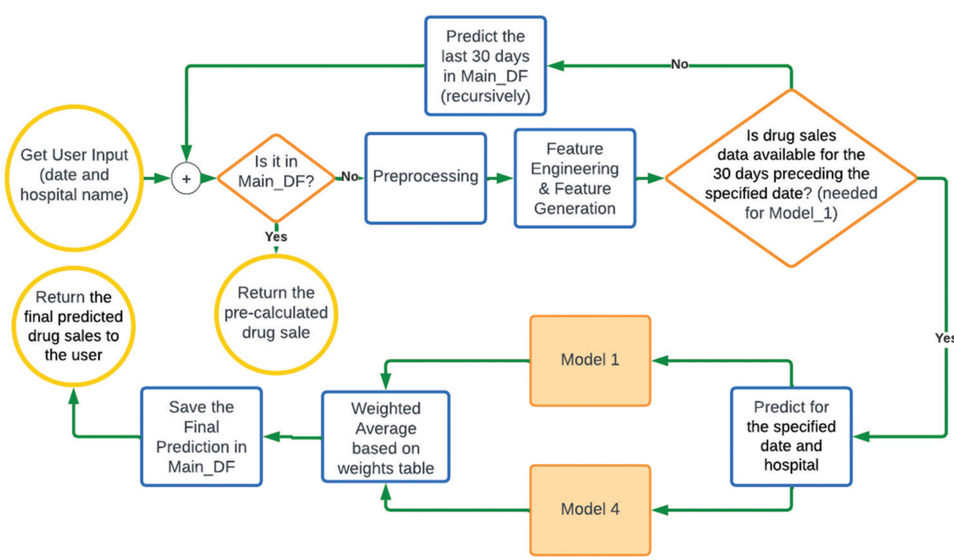

Figure 2: Schema of 5+1 model

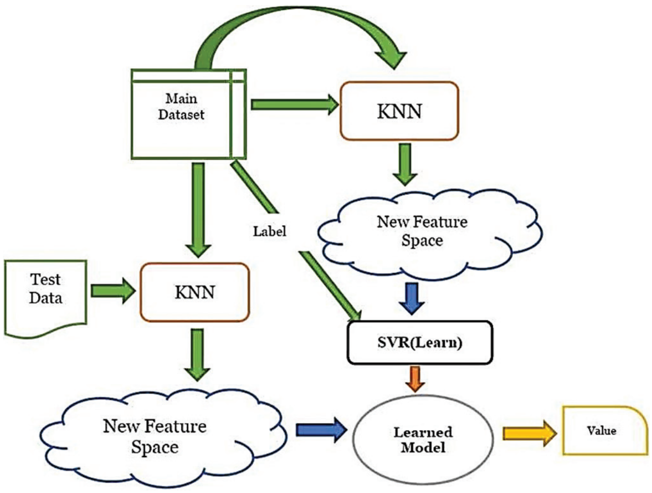

Figure 3: Schema of Ramona pars model

Ranking of competition results

The results of the competition are assessed in two stages. Initially, participants submit their models to the primary drug reviewers for evaluation. The reviewers assess proposed models using preliminary test data. Subsequently, the winning teams progress to the second stage. In each stage, two groups of judges select the winning team based on the following criteria:

1. Innovation in proposed methods for drug demand forecasting
2. Presentation quality in the final report, and articulation of the problem statement
3. The lowest mean squared error (MSE) rate of both stages of the competition.

Performance metrics

The performance metric utilized in this study is MSE. MSE measures the average squared difference between actual and predicted values, serving as a commonly used evaluation tool for predictive model performance, with lower values indicating greater accuracy. The MSE Eq. is as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(d_i - \hat{d}_i)^2$$

Where n is the number of data points, d_i is the observed volume of the demand and $\hat{d}_i$ is the forecasted value.

In the rest of the article, we first provide a detailed description of the competition dataset. Then, the overview of the winning teams and their proposed methods are discussed. Finally, we conclude the article by summarizing the outcomes of the competition and discussing potential avenues for future work.

Description of the Dataset

The datasets used in this study consist of real data records obtained from the Hamedan University of Medical Sciences (HUMS). The HUMS dataset includes 14,860 drug sales records collected between 2019 and 2023 at the university, sourced from 12 pharmacies, as detailed in Table 1. The data is segmented into eight categories: *Pharmacy ID* (a unique identifier for each drug), *WeekDay* (day of the week), *SaleDate* (date of drug sale), *SaleCount* (daily turnover), *Customers* (number of customers per day), *Open* (pharmacy status-open[1] or closed [0]), *Discount* (price reduction on pharmaceutical products), *StateHoliday* (state holiday indicator, i.e., 1 for yes, 0 for no), and *SchoolHoliday* (school holiday indicator, i.e., 1 for yes, 0 for no).

The overall dataset was divided into three categories: training, testing, and real-time testing, as reported in Table 2. The training data were provided to the teams along with the corresponding labels. The real-time testing dataset is intended for model evaluation during the initial and final phases of the competition, and it remained unpublished until the day of the competition.

The Outcome of the Competition

In the first stage of the competition, four teams, namely, *5+1*, *Plexus*, *Salamat Fanavaran*, and *Ramona Pars*, submitted their models after a 6-month working duration on the train data. The proposed models were then evaluated using the first stage test dataset based on MSE, with the results presented in Figure 1. Results are depicted in logarithmic scale for better comparable appearance. As can be revealed from Figure 1, *5+1* and the *Ramona Pars* models achieved the best results, respectively. The high performance of these models can be attributed to effective data preprocessing and the use of various data enrichment methods. Consequently, the models from *5+1* and *Ramona Pars* have been invited to participate in the final stage of the competition for further evaluation.

The final stage of the competition was held live and the final test data were made available to the participants, as described in Table 2. In the following, the proposed methods of the best two teams that reached the final evaluation are described briefly.

5+1 team description

5+1 team introduces a hybrid machine learning and time series model to forecast the demand as shown in Figure 2. Utilizing calendar, climatic features, and historical demand data, the model improves the accuracy of the drug demand predictions. In the initial phase, data dimension expansion transforms four initial features into 19 final features, incorporating relevant climatic and temporal factors identified in previous research as described in Table 3. Subsequently, a sophisticated model utilizing LSTM networks is employed. The proposed ensemble model calculates outcomes from two primary models, combining hospital consumption figures through weighted averages. One primary model, incorporating weather data and past drug consumption, shows commendable

Table 1: Description of the Hamedan University of Medical Sciences dataset used for the competition

Feature	Type date
Pharmacy ID	The numeral of between 1 and 12
WeekDay	The numeral of between 1 and 7
SaleDate	Date between 2019 and 2022
SaleCount	Numeral between 0 and 557
Customers	Numeral between 1 and 10
Open	Binary 0 and 1
Discount	Binary 0 and 1
SchoolHoliday	Binary 0 and 1
StateHoliday	Binary 0 and 1

Table 2: Dataset composition in each stage of the competition

Dataset	Range	Count
Training	January 05, 2019–December 31, 2022	14,290
Testing	January 01, 2023–October 31, 2023	285
Real-time testing	November 1, 2023–November 30, 2023	285

Table 3: Final features used by 5+1 for prediction

Feature name	Description	Data type
Pharmacy ID*	Pharmacy index	Integer
SaleDate*	Drug sale date	Date
WeekDay*	Weekly days index	Integer
SaleCount*	Drug order quantity	Integer
SchoolHolidays*	Closure=1 Nonclosure=0	Boolean
COVID-19 state	COVID-19 days=1, peak of COVID=2, non-COVID=0	Integer
Winter season	Cold=1 and warm=0	Boolean
DoY	Day of the year from start	Integer
DoM	Day of the month from start	Integer
Max_temp	Maximum day temperature	Integer
Min_temp	Minimum day temperature	Integer
Temp	Average day temperature	Integer
RH	Day humidity	Float
Precip	Day precipitation	Float
Wind_spd	Day wind speed	Float
Solar_rad	Day solar radiation	Float
Snow	Snowfall amount	Float

Original features are identified by *

short-term accuracy. However, iterative calculation of long-term drug consumption introduces a potential for increased errors. Conversely, a second model, relying solely on weather data, exhibits superior long-term accuracy. The ensemble model outperforms traditional demand prediction models, which demonstrates the ability to predict both short-term and long-term scenarios by combining various ensemble models. By leveraging model strengths, it empowers hospitals to manage drug inventory effectively, addressing immediate and prolonged demands. This multifaceted approach enhances overall supply chain efficiency, enabling informed decisions for optimal resource allocation and operational performance in healthcare institutions.

Ramona pars team description

Ramona Pars employs a combination of K-Nearest Neighbors (KNN) and support vector regression (SVR) methods for drug demand forecasting, which enhances demand accuracy by transforming the denomination features as shown in Figure 3. Initially, the date feature is decomposed into three components: day, month, and year, resulting in the creation of three new features. Subsequently, utilizing an algorithm analogous to the KNN algorithm, these features are mapped to ten new features derived from the initial dataset, generating a new feature space. A linear SVR model is then trained using these newly created features. Finally, during the testing phase, the same procedure is applied to each new dataset, and the trained model is utilized to predict the output values for the desired test data.

Conclusion

This article looks at models presented in forecasting the drug sales challenge of IAI-2023 in a real dataset from Hamedan University of Medical Science. The proposed models focus on machine learning approaches employing various innovative methods for enhancing the accuracy of drug demand forecasting. Looking at all received results shows all achieved MSE results are upper than 1, which may be due to the noise in the dataset and lack of data records.

Looking at all the winning algorithms of IAI-2023 reveals several interesting findings. (1) All prediction methods are based on machine learning algorithms. The most popular methods are KNN, SVR, RNN, and LSTM. (2) All the methods put a major part of their focus on feature engineering to increase the effective information and hence the resulting accuracy. For example, one team added new dimensions of climatic and calendar features and another team changed the data dimension using the KNN approach.

Financial support and sponsorship

This work was supported by the Isfahan Elite Foundation (IEF), which sponsored the Isfahan Artificial Intelligence Event 2023 (IAI2023). The IEF organized the event and provided financial support for the 10 challenges, including Challenge VIII: Drug Demand Forecasting. Several winners received prizes from the IEF.

Conflicts of interest

The authors declare the following potential conflicts of interest:

- ZZ and AM were the organizers of the Isfahan AI (IAI) 2023 competitions on behalf of Isfahan Elite Foundation (IEF), which included 10 challenges.

- MJ and MP served as scientific committee members for Challenge VIII: Drug Demand Forecasting. They were responsible for evaluating the methodologies and results of all participant teams.
- RR, AG, AAS, HB, PS, BSH, ARN, FM, and SJ are members of the winning teams in this challenge.

None of the organizers and scientific committee members (MJ, MP, ZZ and AM) contributed to the development of the methods used by the participating teams. The final decision regarding the winners was made by the policy council members based on the following criteria:

- Technical contribution in developed models by teams,
- The results on initial and final test data of each team,
- The submitted reports and teams' presentations.

The authors have disclosed these relationships to ensure transparency and maintain the integrity of the research.

Acknowledgements

We extend our appreciation to the policy council members, Prof. Behrouz Minaei Bidgoli from Iran University of Science and Technology, Prof. Mohammad Hassan Moradi from Amirkabir University of Technology, Prof. Heshaam Faili from University of Tehran, Prof. Emad Fatemizadeh from Sharif University of Technology, and Prof. Hossein Rabbani from Isfahan University of Medical Sciences, for making the final decisions regarding the winners of all 10 challenges.

Finally, we thank the anonymous reviewers for their valuable feedback and contributions to the improvement of this paper.

References

1. Bown CP, Bollyky TJ. How COVID-19 vaccine supply chains emerged in the midst of a pandemic. World Econ 2022;45:468-522.
2. Wouters OJ, Shadlen KC, Salcher-Konrad M, Pollard AJ, Larson HJ, Teerawattananon Y, *et al.* Challenges in ensuring global access to COVID-19 vaccines: Production, affordability, allocation, and deployment. Lancet 2021;397:1023-34.
3. Hoeltgebaum H, Borenstein D, Fernandes C, Veiga Á. A score-driven model of short-term demand forecasting for retail distribution centers. J Retail 2021;97:715-25.
4. Verstraete G, Aghezzaf EH, Desmet B. A leading macroeconomic indicators based framework to automatically generate tactical sales forecasts. Comput Ind Eng 2020;139:106169.
5. Merkuryeva G, Valberga A, Smirnov A. Demand forecasting in pharmaceutical supply chains: A case study. Procedia Comput Sci 2019;149:3-10.
6. Schmidt A, Kabir MW, Hoque MT. Machine learning based restaurant sales forecasting. Mach Learn Knowl Extraction 2022;4:105-30.
7. Khosravi H, Thaker AH, Donovan J, Ranade V, Unnikrishnan S. Artificial intelligence and classic methods to segment and characterize spherical objects in micrographs of industrial emulsions. Int J Pharm 2024;649:123633.
8. Sousa RM, Hannachi S, Ramos GN. Statistical and deep learning models for forecasting drug distribution in the brazilian public health system. In 2019 8th Brazilian Conference on Intelligent Systems (BRACIS). IEEE; 2019. p. 723-8.
9. Zwaida TA, Pham C, Beauregard Y. Optimization of inventory management to prevent drug shortages in the hospital supply chain. Appl Sci 2021;11:2726.
10. Abbas K, Afaq M, Khan TA, Song WC. A blockchain and machine learning-based drug supply chain management and recommendation system for smart pharmaceutical industry. Electronics 2020;9:1-31.
11. Syahrir I, Suparno S, Vanany I. A Proposed Model for Drug Demand Forecasting and Ordering Inventory System for Dengue Endemic. Oper Supply Chain Manage 2022;15:69-78.