International Journal of Data Science and Analytics
https://doi.org/10.1007/s41060-024-00530-z

REVIEW


# Predicting the pharmaceutical needs of hospitals using machine learning algorithms

**Amir Hossein Nabizadeh**[1,2] · **Mohammad Mehdi Ghaemi**[1] · **Daniel Goncalves**[2,3]

Received: 1 September 2023 / Accepted: 26 February 2024


**Abstract**
People's lives are always threatened by various diseases. The role of health and medical services, in particular medicine, is undeniable in protecting their lives. Timely preparation and providing medicine for patients is vital since medicine shortage can endanger their lives while excessive accumulation of medicine can put them at expiration risk and waste health budgets. To this end, in this paper, we aim at introducing a model for the prediction of commonly used medicine (type and amount) in hospitals. For that, in our applied research, we initially used patients' data from the Afzalipur Hospital in Kerman collected for 3 years consisting of 283 features, which included over 9351 different medicine and 121,690 patients. Then, nine features were selected using experts' feedback and were fed into the random forest and neural network algorithms. For the prediction task, medicine types and their amounts were predicted for each individual using different training sets. In addition, the right prediction time was also found which is when predictions have a promising accuracy while the executive team of a hospital has enough time to provide the right amounts of the most used medicine. The performance of algorithms was evaluated using a confusion matrix (precision, recall, F1, and accuracy metrics). Our results showed that the random forest had a promising performance in predicting the amounts of the most used medicine for a month using 2 years of data (accuracy 83.3%) while its accuracy in predicting medicine was 35.9%. Therefore, we conclude that the random forest algorithm has the potential to effectively and accurately predict medicine amounts by analyzing large amounts of data and detecting patterns that may not be easily discernible to humans.

**Keywords** Drug target predictions · Random forests · Computer neural network · Machine learning · AI (artificial intelligence)

## 1 Introduction

People's health is completely dependent on medicine, and one of the most important criteria in the health system of any country is the amount of medicine consumed in that country. Medicine often poses an excessive cost to individuals and the healthcare system. Global budget spending on medicine in 2020 was around 1.3 trillion dollars while the United States alone spent around 350 billion dollars on it. These budgets are expected to rise at a rate of 3% to 6% annually worldwide. Funding medicine becomes more challenging when we know that the lack of resources in providing medicine and unequal access to it can negatively influence this rate [1].

Regarding the mentioned challenges, accurate management of supplying medicine is essential. The final goal of this management is the timely preparation and distribution of medicine among patients since the medicine shortage can endanger their lives, and on the other hand, excessive accumulation of medicine can put medicine at risk of expiration and waste health budgets.

In the past, traditional and simple methods have been used for estimating and predicting the required medicine. For instance, providing and storing a huge bulk of flu medicine in the autumn and winter seasons. Today, accurate and

Communicated by Ana Carolina Lorena.

✉ Mohammad Mehdi Ghaemi
m.ghaemi@kmu.ac.ir

Amir Hossein Nabizadeh
amir.nabizadeh@tecnico.ulisboa.pt

Daniel Goncalves
daniel.goncalves@tecnico.ulisboa.pt

[1] Medical Informatics Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman 7616913555, Iran

[2] INESC-ID, Lisboa, Portugal

[3] Instituto Superior Tecnico, Universidade de Lisboa, Lisboa, Portugal

Published online: 15 April 2024

cost-effective methods have been introduced and used in developed countries for the prediction of required medicine that work based on evidence (e.g., dosage) and uncertainty variables [1–4]. These methods focused on various topics, such as medicine prediction [5–11], or medicine cost prediction [12–17].

Mahajan and Kumar implemented a time-series model to predict the sales of pharmaceutical distribution companies [18] using various approaches, such as the ARIMA method [19], neural networks [20], advanced neural networks [21], and fuzzy neural networks [22]. Their results showed that the method could make an approximate prediction in real-time. Neural networks have also been used in [23]. In this study, Chang and his colleagues performed a weekly forecasting of pharmaceutical products using a deep neural network [24]. For the evaluation, a week-ahead sales prediction was done using 3 years of daily sales, which were grouped weekly. In another study, researchers used quantum neural networks-QNNs [25] to predict the exact generic medicine for a patient by considering disease factors [26]. For that, initially, the symptoms of disease (105 symptoms) were categorized by experts. To assess, the results of support vector machines (SVMs) [27], Bayes [28], random forest—RF [29], and QNN were compared, which the QNN model had reached the accuracy of 95%.

Alves predicted the demand for hospital consumables [30] using neural networks—RNN [31], SVM, and RF with the help of parametric methods, such as Holt-Winters model exponential smoothing [32], and the ARIMA method [19]. Results revealed that the machine learning algorithms were able to produce the lowest average prediction error. It was also concluded that the classification of items directly affects the performance of the prediction models. In a similar study, authors presented a model to use large data sets (50 variables collected from 101,766 patients in 120 hospitals) for recommending new medicine to diabetic patients [33]. Ten variables were selected via feature selection methods and fed into the RF, multilayer perceptron—MLP [34], logistic regression—LR [35], J48 [36], and SVM algorithms, which RF and MLP had the best performance. RF, NN, and other methods like linear regression were also applied by Mbonyinshuti et al. to predict the future trends in demanding the ten most used medicine in Rwanda [37]. To evaluate the results, predictions were compared with the actual values. Results showed that the RF had the best performance. In another study, the same authors used the RF model to predict the demand for essential medicine for non-communicable diseases (NCD) based on past consumption data [38]. Their data included more than 500 medicine, and only 17 of the most used ones were selected for the prediction task. The evaluation showed that the RF model could accurately predict the medicine for a month.

In [39], a model was introduced to assist patients in finding appropriate medicine for their diseases, which can be extended by hospitals to predict their needs. This model analyzed the history of a patient to check for any side effects of a suggested medicine. In addition, it checked the weather conditions and local maps (via Google's API) where the patient was located, so that he/she could find nearby drugstores that have the medicine. In another study, researchers designed a preliminary prediction approach. It worked based on a dynamic meta-analysis to predict the temporary sale of a new generic medicine when there is a complete lack of information [40]. During the evaluation, a dynamic meta-analysis of the release parameters was able to predict the cycle of the medicine. Performance evaluation verified a high degree of accuracy between the previously observed values and its predicted average.

Despite the use of modern predicting methods by the aforementioned studies, these studies often ignored the time factor in their predictions. In addition, these methods were usually implemented and used in developed countries and have not been widespread in third-world countries. In the least developed countries, hospitals and medical centers still use traditional methods to estimate their required medicine. Traditional methods are not very complicated and do not have high accuracy [1]. The low accuracy of these methods can have various reasons, including not taking into account conditions, such as the amount of medicine produced by companies, medicine price variation, lack of raw materials, etc., which can have a high and direct impact on the amount of production and consumption of medicine. Machine learning (ML) methods can be applied as a solution to prevent an unreasonable increase in medicine stock and, as a result, mutation of their price.

To this end, in this study, our goal is to apply ML-based methods to predict the need (amounts and types of medicine) for the most used medicine in a hospital and evaluate the prediction results. To achieve our goal, we split our main goal into sub-goals and achieve them one by one. Our subgoals are as follows:

- Finding the most used medicine in a hospital.
- Detecting informative data features for the prediction task.
- Developing ML algorithms to predict amounts and types of medicine.
- Evaluating the performance of the prediction algorithms.
- Improving the accuracy of the prediction algorithms.

To achieve our goal, we initially collected data on medicine used by patients in Afzalipur Hospital in Kerman. It was collected from 2018 to 2020 and included the data of 9351 different medicine prescribed to 121,690 patients.

In brief, the contribution of our work is the following:

- Generating a hospital dataset collected for 3 years including 283 features.
- Proposing a model to predict the most used medicine of a hospital using ML techniques a month ahead.
- Proposing a model to predict the amounts of the most used medicine of a hospital using ML techniques a month ahead.
- Finding prediction time. It is when the predictions have a promising accuracy while management team of a hospital has enough time to provide the right amounts of the most used medicine.
- Finding a prediction model that constantly has a promising accuracy.
- Enhancing the prediction accuracy using a class balancing technique.

## 2 Research methodology

Our study is an applied research. In this section, we detail our prediction approach, describe what type of data is collected, how the most used medicine are identified, and what algorithms are used for the prediction task. In addition, we explain how we apply a feature selection technique to drop insignificant data attributes to improve prediction accuracy. Figure 1 depicts a general view of our approach and the steps it takes for the prediction.

### 2.1 Data collection

Our data were collected from the Afzalipur Hospital in Kerman from the beginning of 2018 to the end of 2020 (the hospital did not give us the permission to collect more than 3 years of data), including 283 clinical and demographic features. This data was collected using Shafa software, which is designed to store prescribed medicine for patients of this hospital. The inclusion and exclusion criteria for data gathering were:

**Inclusion criteria:**

- Inpatients and Emergency Department (ED) patients' data including:
  - Demographic data, such as age and gender.
  - Clinical data, such as laboratory tests, prescribed medicine, diagnosis codes, physicians codes.
  - Geographic data, such as patients' home address.

**Exclusion criteria:**

- Outpatient data.
- Irrelevant patients' data, such as phone number, and education.
- Insurance and payment related data.

Table 1 details the data for each year. As presented in this table, the number of admission in 2020 was much less than in 2018 and 19. It was due to the reason that this hospital was mainly assigned for the Coronavirus disease 2019 (COVID-19) cases.

Figure 2 shows the prescription amounts of each medicine. The most used medicine was prescribed less than 400 K times in 3 years, and the least one was cisplatin/vial/0.5 mg/ml, 100 ml, which was prescribed only once. Among 9351 medicine, 1057 of them have been prescribed more than 1000 times. After an initial analysis, it was found that the medicine with the highest prescription was "plastic distilled water," which is not a medicine. So, these kinds of medical items were ignored in our data.

### 2.2 Data pre-process

Preprocessing played a pivotal role in our process and took a considerable amount of our time. Our preprocessing phase was conducted in three steps using R programming language. In the first step, we ignored the records having missing (null) or invalid data and did not consider them in the prediction process.

In the second step, duplicated variables were removed from the data set. As mentioned earlier, the initial data consisted of 283 variables, some duplicated and stored with different names in different columns. Therefore, these types of variables were also ignored.

In the last step, it was found that several admissions had negative values for the amounts of prescribed medicine. It happened due to mistyping the amounts by operators, and also the improper design of the system in not allowing invalid amounts. After a consultation with an expert, these values were considered positive.

### 2.3 Identify target medicine

As stated before, we aim at predicting the type and amounts of the most used medicine in a hospital. For that, we initially needed to estimate the amounts of each medicine. Due to the negligence of operators and faulty design of the database, there were medicine stored having different IDs (Table 2). For instance, chlorsodium dextrose serum was stored having five various IDs. To solve this problem, we unified all IDs for a single medicine and accumulated their amounts. Due to the sensitivity of our field, this step was done manually. We then analyzed the amounts of all medicine and found that 20 medicine had a clear distance from the rest. Therefore, we only concentrated on them. Table 2 shows the top 20 medicine along with their ID and amounts.

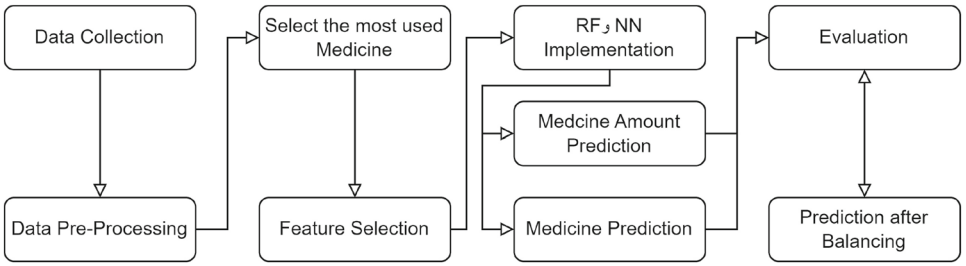



**Fig. 1** Research methodology

**Table 1** Number of admissions, medicine, and total amounts of prescription

| Year | # Admission | # Different medicine | # Prescribed medicine | # System records |
|---|---|---|---|---|
| 2018 | 48,012 | 4032 | 4,063,539 | 2,132,075 |
| 2019 | 45,171 | 4066 | 3,994,663 | 2,231,849 |
| 2020 | 28,507 | 3467 | 42,35,229 | 2,029,816 |
| Total | 121,690 | 9351 | 1,2293,431 | 6,393,740 |

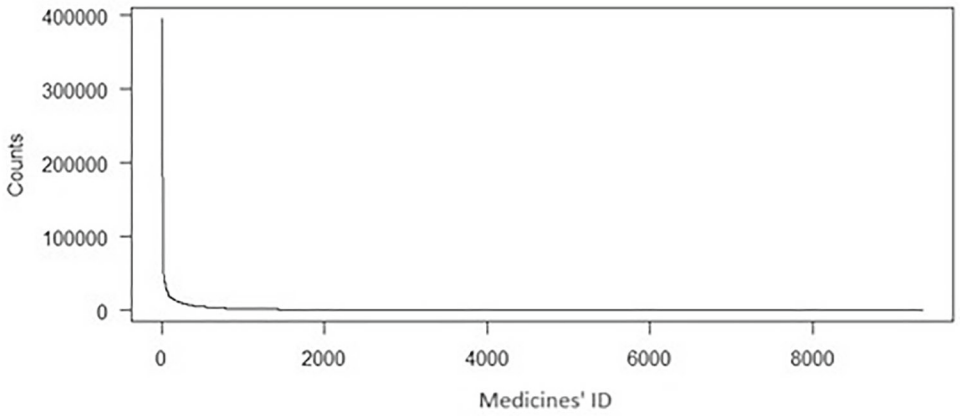



**Fig. 2** Medicine and their amounts

## 2.4 Feature selection

Feature selection is a technique to identify the relevant and informative data attributes while discarding the redundant and irrelevant ones [41, 42]. It has several benefits. It avoids an overfitting problem [43], facilitates the interpretability (easy to understand) of the prediction results [44–46], enhances the learning speed, reduces the storage volume, and decreases the noise caused by redundant and irrelevant attributes [47]. To this end, we used Boruta method that works based on an RF algorithm [48]. It iteratively compares the importance of original attributes with the shadow ones (created by shuffling the original attributes). The original attributes having lower importance than the shadow ones are dropped while the higher ones are kept as the main (confirmed) attributes. The shadow attributes are regenerated in each iteration. Boruta stops when no confirmed attributes are left or the stop criteria is met (i.e., max iteration).

By applying the Boruta, only one attribute was detected as the confirmed one. After discussing the results with experts, we concluded that we do the feature selection manually using the experts' feedback. Therefore, after discarding the duplicated features (40 features), constant features (100 features), and the irrelevant ones (134 features), we ended up with nine features including patients' age, blood type, sex, diagnosis code, disease code, physician code, admission month, prescribed medicine and their amounts for each patient.

## 2.5 Prediction algorithms

After determining the informative features and discarding the irrelevant ones, we have applied the regular form of two predictive algorithms, random forest (RF) and neural network (NN) for the prediction task. The hyper-parameters of these models were tuned using a fivefold cross-validation technique (detailed in Sect. 2.6).

Random forest (RF) is a combination of decision tree predictors. Trees are depending on values of random vectors, which are sampled independently while having the same distribution [49, 50]. RF is commonly used by researchers and has several advantages, such as high accuracy, modeling the complex interactions among predictor attributes, handling data with numerous features, working well with both continuous and categorical values, being flexible to both regression

**Table 2** Top 20 medicine with their ID and amounts

| Medicine | Amount | ID01 | ID02 | ID03 | ID04 | ID05 | ID06 |
|---|---|---|---|---|---|---|---|
| Dextrose chlorsodium serum 0.3% injection 0.5 l | 178,025 | 65,165 | 53,607 | 54,922 | 65,031 | 52,363 | – |
| Sodium chloride serum injection bottle/0.9%/0.5 l | 177,304 | 42,870 | 52,384 | 54,345 | 46,789 | 65,004 | 56,707 |
| Sodium heparin/ampoule/5000 units/1 l | 128,955 | 51,857 | 55,200 | 53,788 | 56,339 | – | – |
| Clindamycin Phosphate/Ampoule/300 mg/2 ml | 89,876 | 51,844 | 46,009 | 55,275 | – | – | – |
| Vancomycin/vial/500 mg | 58,411 | 56,543 | 48,179 | – | – | – | – |
| Dexamethasone/ampoule/8 mg/2 ml | 57,110 | 65,019 | 51,847 | – | – | – | – |
| Pantazol vial / 40 mg | 50,976 | 47,138 | – | – | – | – | – |
| Ceftriaxone/vial/1 gram | 50,264 | 45,785 | 51,841 | – | – | – | – |
| Calcium gluconate/ ampoule/ 10% | 31,634 | 56,200 | – | – | – | – | – |
| Potassium chloride/ampoule/10 ml | 29,935 | 57,379 | – | – | – | – | – |
| Naloxone/ampoule/0.4mg/1ml | 29,352 | 56,354 | – | – | – | – | – |
| Azithromycin/capsule/250 mg | 27,688 | 56,221 | – | – | – | – | – |
| Calcium carbonate/pill/500 mg | 27,028 | 56,178 | – | – | – | – | – |
| Pantoprazole/pill/40 mg | 26,972 | 43,079 | – | – | – | – | – |
| Sodium chloride serum bottle/injection/0.4%/0.5 liter | 26,117 | 51,921 | – | – | – | – | – |
| Famotidine/pill/40 mg | 22,600 | 64,620 | – | – | – | – | – |
| Furosemide/ampoule/20 mg/2 ml | 22,187 | 46,813 | – | – | – | – | – |
| Prednisolone/pill/5 mg | 22,026 | 47,310 | – | – | – | – | – |
| Ciprofloxacin/bag/200 mg/100 ml | 21,251 | 46,524 | – | – | – | – | – |
| Ranitidine/ampoule/50 mg | 19,657 | 46,831 | – | – | – | – | – |

Sorted by their amount

and classification problems, reducing the over-fitting issue, parallel computation, and robustness to emissions data due to random sampling [51–53]. Neural networks (NNs) are made of node layers, containing one input layer, one or more hidden layers, and one output layer. Each node (neuron) is connected to another one and has an associated weight and threshold. If the output of a node is more than a specified threshold, that node is activated, sending data to the next layer. NN provide various advantages over traditional algorithms. They learn from data and can implicitly detect complex nonlinear relationships between dependent and independent variables and also detect all possible interactions between predictor variables. Therefore, they can be applied to solve complex problems and are able to generalize (i.e., recognize patterns in data that traditional algorithms may not). Moreover, they are scalable and can handle large amounts of data quickly and accurately. [54–58]. Due to the mentioned advantages, RF and NN algorithms were selected for the prediction tasks.

Our prediction was twofold. We initially predicted the amounts of medicine. Due to the reason that we had a limited amount of data for each medicine and the amounts of medicine varied a lot, therefore, prediction of exact amounts of a medicine could not be achievable since having one or a few examples for each possible value does not make a pattern. Therefore, by converting the amounts of medicine into intervals, we would have at least X points for each medicine which helps ML models to perform more accurately. Hence, all amounts of the prescribed medicine for patients were converted into 10-point intervals (e.g., all amounts from 1 to 10 were assigned to the first interval and so forth), and then these intervals were predicted. We then predicted the type of medicine.

In total, twenty-four models were built for the prediction task:

- Twelve predictive models (medicine and their amounts) for a **month** using RF and NN methods using 1, 2, and 3 years of data (four models for each year).
- Twelve predictive models (medicine and their amounts) for a **season** using RF and NN methods using 1, 2, and 3 years of data (four models for each year).

In our study, our training sets were the data of 1, 2, and 3 years while the test set was the data of a month/season. Different amounts of data were used for training to analyze how much data is needed for having a promising prediction accuracy, and also monitor which method consistently works better.

### 2.6 Evaluation

Prescribing the right medicine is vital since the wrong prescription can have the risk of death. Therefore, the high accuracy of our models is of special importance. Our evaluation was twofold. We initially found the optimal values for the hyper-parameters of the models. For that, we used a five-fold cross-validation technique along with the TuneLength method in the R programming language. TuneLength considered five default values for each parameter and selected a value that provided the highest accuracy for the prediction. Then, we used the confusion matrix to estimate the performance of the models. Based on this matrix, we calculated various measures, such as accuracy, precision, recall, and $F$-1 (Eqs. 1 to 4). The results for the last three measures (recall, precision, and $F$-1) are presented in A.3.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$F1 = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

In the above equations, the true positive (TP), false positive (FP), false positive (FP), and true negative (TN), are as follows:

- True positive (TP): These are the cases where the model predicted the positive class correctly.
- False positive (FP): These are the cases where the model predicted the positive class incorrectly.
- False negative (FN): These are the cases where the model predicted the negative class incorrectly
- True negative (TN): These are the cases where the model predicted the negative class correctly.

## 3 Results

As explained, our predictions were made for a month/season using different sets of training data:

- Prediction of medicine and their amounts for a **month** using data of 1/2/3 years.

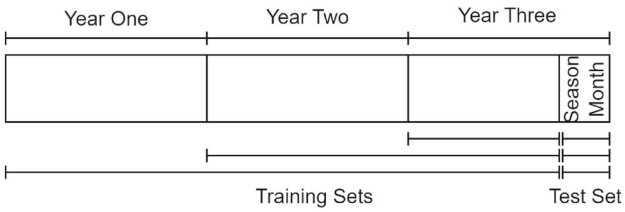



**Fig. 3** Train and test sets

**Table 3** Train and test sizes (number of records)

| | Predict a month | | |
|---|---|---|---|
| Training set | 1 year | 2 years | 3 years |
| Train size | 100,267 | 183,645 | 269,673 |
| Test size | 9050 | 9050 | 9050 |
| | Predict a season | | |
| Training set | 1 year | 2 years | 3 years |
| Train size | 80,103 | 163,481 | 249,509 |
| Test size | 20,164 | 20,164 | 20,164 |

- Prediction of medicine and their amounts for a **season** using data of 1/2/3 years.

We initially evaluated the prediction accuracy of NN and RF models. To generate the models, the data of one month/season were used as a test set, and 1/2/3 years of data were used as training sets (Fig. 3). In our study, we predict the future requirements of the hospital based on the current and past data. Therefore, since our prediction case is a time-series case, we should have a sort of control over generating test sets to ensure that the training sets are anterior to test sets to have more reliable results. The sizes of these train and test sets are mentioned in Table 3. As detailed in Table 1, our general dataset included over 121,690 patients (described in Sect. 2.1) and 9351 different medicine, which was collected over 3 years (2018–2020).

To comprehend the results in this section better, we graphically illustrate the frequencies of genders through months (see Fig. 4). Figure 4 reveals that the frequencies of males and females are relatively similar across all months.

Based on the results presented in Table 4, RF often outperformed the NN models. In this table, the accuracy using 1 year of training data is usually lower than the rest of the years. It could be due to not having enough data for training the models. On the other hand, the accuracy of results using 3 years of data is less than 2 years, which might be because of increasing the noise in the collected data that negatively influences the accuracy. Considering the results in Table 4, both models had their best overall performance using 2 years of training data.

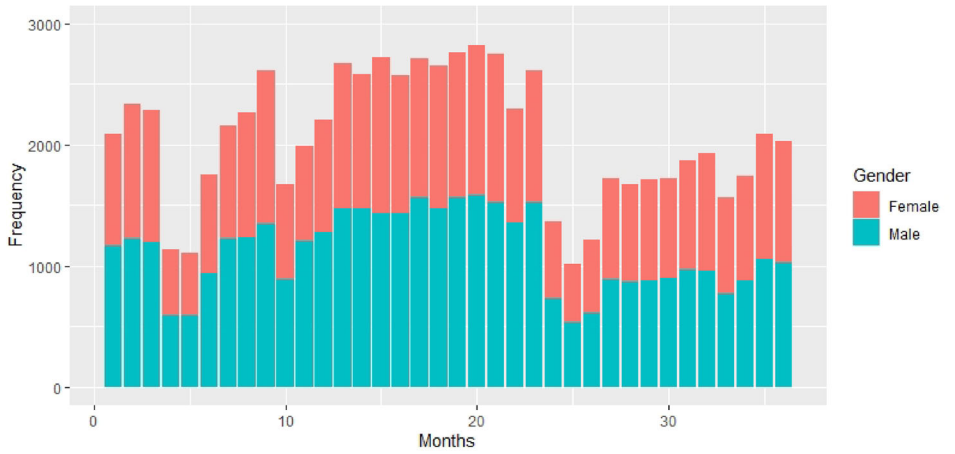



**Fig. 4** Gender frequencies per month (2018–2020)

**Table 4** RF and NN prediction accuracy (before balance)

| Test sets | A month | | | A season | | | |
|---|---|---|---|---|---|---|---|
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | Medicine |
| RF | **24.22%** | **31.89%** | **27.15%** | **28.06%** | **28.49%** | 28.25% | |
| NN | 8.11% | 28.79% | 21.20% | 9.23% | 22.70% | **28.90%** | |
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | Amounts |
| RF | **77.06%** | **83.35%** | **81.55%** | 77.69% | 81.33% | NA | |
| NN | 77.06% | 10.77% | 81.55% | **77.74%** | **81.36%** | **79.31%** | |

The best results are bolded

By analyzing the prediction results, we realized that the accuracy of predicting medicine is significantly lower than the accuracy of their amounts. It could be due to the reason that all medicine were not prescribed in a balanced way, and the prescription of some medicine was far more than the others (Fig. 5). Since some predictive models are very sensitive to imbalanced data, it is possible that this issue negatively influences their performance. For this purpose, after balancing the data of the prescribed medicine via SMOTE technique [59], we re-evaluated the accuracy of the models. As presented in Table 5, the accuracy of predicting medicine has been significantly improved.

Although the accuracy of medicine prediction has improved after balancing, the accuracy of their amounts, which was acceptable before balancing, has decreased. To this end, we decided to use balancing only for the prediction of medicine and not use it for their amounts.

Due to the reason that RF often outperformed the NN models and normally, the higher accuracies were obtained using 2 years of training data, our selection for the prediction is the RF model using 2 years of training data for the prediction of a month. The architecture for the final NN and RF models is presented in Table 6. In addition, the error rates of NN and RF models in Fig. 6 indicate the the error rates of the NN and RF models using 2 years of data for predicting a month are shown in Fig. 6. As presented in this figure, the error rates for both models are very similar.

The confusion matrices of the models in Fig. 6 are presented in Tables 7, 8, 9 and 10. As shown in these tables, the error in the prediction of medicines' amounts is much less than in predicting the medicine. These results are compatible with the results presented in Table 4, where the accuracy for the prediction of medicines' amounts is higher than the prediction of medicine.

## 4 Discussion

The evaluation results showed both RF and NN models had almost the same accuracies (RF model had a slightly higher accuracy) while the time to generate RF models was far more than the NN models (results in A.2). The highest accuracy in predicting the amount of medicine belonged to the RF model, 83.3% (for a month) and 81.3% (for a season) using 2 years of data. Medicine prediction was far less accurate, and the best results using the RF model for a month and a season were 31.8% and 28.4%, respectively. Although the prediction accuracy for the amount of medicine was acceptable, it was low in medicine prediction. To enhance the accuracy, we balanced the prescribed medicine classes (reproduced medicines that were prescribed much less than the others). After balancing, the accuracies for medicine prediction raised to 35.9% (a month) and 40.3% (a season).

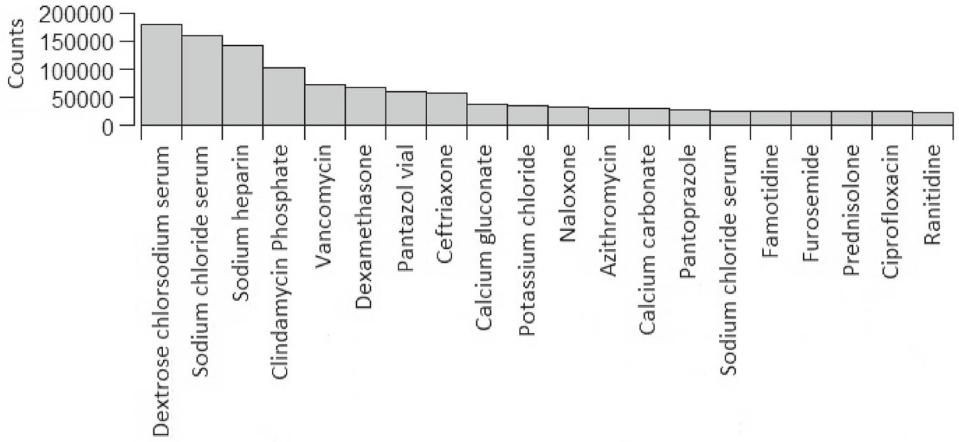



**Fig. 5** Top 20 medicine and their amounts

**Table 5** RF and NN prediction accuracy (after balance)

| Test sets | A month | | | A season | | | |
|---|---|---|---|---|---|---|---|
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | Medicine |
| RF | **83.18%** | **35.95%** | **39.53%** | **91.96%** | **40.36%** | **43.21%** | |
| NN | 8.09% | 28.76% | 29.33% | 9.22% | 22.68% | 28.90% | |
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | Amounts |
| RF | **64.27%** | 81.71% | **77.09%** | **72.53%** | **78.57%** | **75.21%** | |
| NN | 64.16% | **81.82%** | 77.09% | 71.90% | 78.57% | 75.10% | |

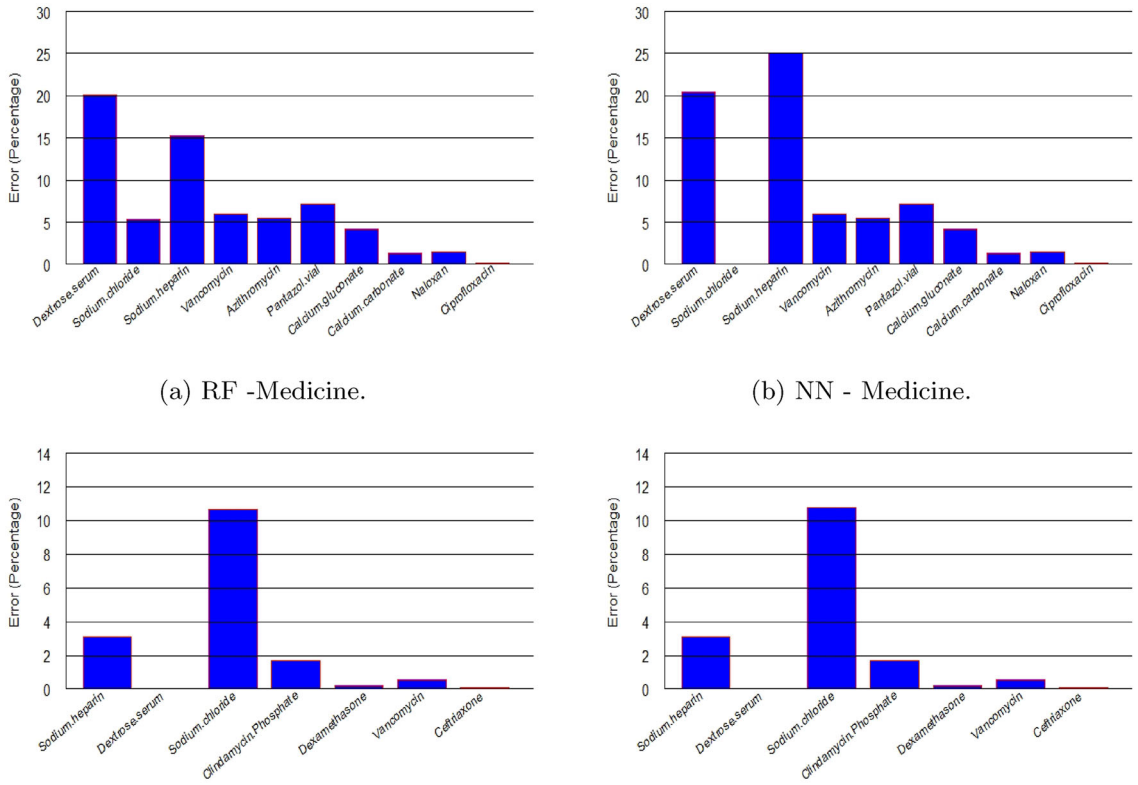
The best results are bolded

**Table 6** Architecture of final models (RF and NN)

| Test set | A month | | A season | | – |
|---|---|---|---|---|---|
| Train set | 2 years | 2 years | 2 years | 2 years | – |
| Prediction | Medicine | Amounts | Medicine | Amounts | – |
| Random forest | 500 | 500 | 500 | 500 | # Trees |
| | 2759 | 1521 | 2799 | 2325 | # Nodes |
| | 17 | 24 | 18 | 28 | # Classes |
| Neural Network | 7–1–17 | 7–1–24 | 7–1–18 | 7–1–28 | Network topology |
| | 42 | 56 | 44 | 64 | # of Weights |
| | 5 | 5 | 5 | 5 | # of repeats |
| | 0.1 | 0.1 | 0.1 | 0.1 | Learning rate |
| | 26 | 33 | 27 | 37 | # Hidden layers |

**Table 7** Confusion matrix of NN model to predict 1-month medicine amount using 2 years of data

| | Predictions | | | | | | |
|---|---|---|---|---|---|---|---|
| | Sh | Ds | Sc | Cp | De | Va | Ce |
| Real values | | | | | | | |
| Sodium heparin (Sh) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Dextrose serum (Ds) | 28 | 753 | 97 | 15 | 2 | 5 | 1 |
| Sodium chloride (Sc) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Clindamycin phosphate (Cp) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Dexamethasone (De) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Vancomycin (Va) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Ceftriaxone (Ce) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |




(a) RF -Medicine.

(b) NN - Medicine.

(c) RF - Medicine Amount.

(d) NN - Medicine Amount.

**Fig. 6** Error rates. Train set: 2 years, test set: 1 month

**Table 8** Confusion matrix of NN model to predict 1-month medicine using 2 years of data

| | Predictions | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Ds | Sc | Sh | Va | Az | Pv | Cg | Cc | Na | Ci |
| Real values | | | | | | | | | | |
| Dextrose serum (Ds) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Sodium chloride (Sc) | 185 | 260 | 226 | 54 | 49 | 64 | 38 | 12 | 13 | 2 |
| Sodium heparin (Sh) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Vancomycin (Va) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Azithromycin (Az) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Pantazol vial (Pv) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Calcium gluconate (Cg) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Calcium carbonate (Cc) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Naloxan (Na) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Ciprofloxacin (Ci) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

It is of interest that our results are compatible with the ones presented in Mbonyinshuti et al. [38]. In this study, the authors used the RF model to predict the need for NCD medicine. Their results showed that the RF could accurately predict the medicine a month ahead. In addition, Iqbal et al. predicted future medicine consumption [60]. They stated that the prediction should be made 15 to 30 days before the actual time of medicine requirement.

## 4.1 Limitations

Like any other study, our study had limitations that can be tackled in the future. These limitations are listed below:

- **Noisy Data**: Results showed that the accuracy did not enhance by increasing the amount of data. It could be because the noise increased with the increase of data, which negatively affects the accuracy.

**Table 9** Confusion matrix of RF model to predict 1-month medicine amount using 2 years of data

| | Predictions | | | | | | |
|---|---|---|---|---|---|---|---|
| | Sh | Ds | Sc | Cp | De | Va | Ce |
| Real values | | | | | | | |
| Sodium heparin (Sh) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Dextrose serum (Ds) | 28 | 753 | 96 | 15 | 2 | 4 | 1 |
| Sodium chloride (Sc) | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| Clindamycin phosphate (Cp) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Dexamethasone (De) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Vancomycin (Va) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Ceftriaxone (Ce) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**Table 10** Confusion matrix of RF model to predict 1-month medicine using 2 years of data

| | Predictions | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Ds | Sc | Sh | Va | Az | Pv | Cg | Cc | Na | Ci |
| Real values | | | | | | | | | | |
| Dextrose serum (Ds) | 3 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Sodium chloride (Sc) | 111 | 212 | 137 | 50 | 42 | 61 | 35 | 10 | 9 | 1 |
| Sodium heparin (Sh) | 71 | 48 | 88 | 4 | 7 | 3 | 3 | 1 | 4 | 1 |
| Vancomycin (Va) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Azithromycin (Az) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Pantazol vial (Pv) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Calcium gluconate (Cg) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Calcium carbonate (Cc) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Naloxan (Na) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Ciprofloxacin (Ci) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

- **Faulty Hospital System**: In the preprocessing phase, we found that the Afzalipur Hospital system has serious problems, including recording similar data in different columns, patient admission with no date, or even recording negative values for the volume of prescribed medicine. Although some of the issues are due to the inaccuracy of the system operator, the appropriate design of the system can diminish the mentioned issues to a large extent. It can result in having certified data, which might enhance the prediction accuracy and guarantee its reliability.
- **Metadata**: We used a series of patient, disease, and hospital variables for our prediction. There are still variables, such as "patient underlying disease" and "medical tests results" that can be explored to enhance the accuracy of predictions.
- **Results Uncertainty**: Despite predicting medicine and their amounts, results are not reliable and cannot be generalized since the data was insufficient, noisy, and was for a limited time frame (only 3 years).
- **Low Accuracy of Medicine Prediction**: The RF model could predict the amount of the most used medicine, but its accuracy for the prediction of medicine was not promising. Although by balancing the data accuracy of predicting medicine improved, these results are still low. Other models that consider time-series data, like ARIMA, might help to raise the accuracy.

### 4.2 Future directions

In this study, we predicted the most used medicine along with their amounts for all departments of a Hospital. Due to variations in the volume of the pharmaceutical needs of different departments, the possibility of accurately predicting these needs was trivial. Therefore, in the future, instead of focusing on all departments, which include a wide variety of prescribed medicine, it is recommended to concentrate only on one department. It can increase prediction accuracy due to the convergence of prescribed medicine.

Although predicting the need for commonly used medicine is of great importance, this prediction could be more beneficial considering other factors, such as medicine price, their criticality, and rarity.

## Conclusions

In this study, we predicted the most used medicine (type and amount) in a hospital. For that, we used a dataset collected

**Table 11** Accuracy of the generated models

| Test sets | A month | | | A season | | | |
|---|---|---|---|---|---|---|---|
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | Medicine |
| RF | **47.15%** | **36.93%** | **34.00%** | **48.29%** | **37.13%** | **34.10%** | |
| NN | 37.76% | 31.76% | 30.40% | 39.73% | 32.08% | 30.43% | |
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | Amounts |
| RF | **83.01%** | 83.59% | **80.65%** | 83.30% | **82.39%** | NA | |
| NN | 82.98% | **83.63%** | 80.65% | **83.30%** | 82.39% | 80.69% | |

The best results are bolded

**Table 12** Time to generate models (s)

| Test sets | A month | | | A season | | | |
|---|---|---|---|---|---|---|---|
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | Medicine |
| RF | 240.95 | 841.95 | 1484.9 | 222.41 | 819.41 | 1455.8 | |
| NN | **32.01** | **110.22** | **190.8** | **29.47** | **104.7** | **191.53** | |
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | Amounts |
| RF | 73.8 | 70.82 | 517.6 | 68.55 | 259.44 | 630.22 | |
| NN | **51.18** | **45.88** | **310.83** | **46.9** | **157.7** | **298.73** | |

The best results are bolded

from a Hospital for 3 years. Nine features of the collected data, including patients' age, blood type, sex, diagnosis code, disease code, physician code, admission month, prescribed medicine, and their amounts were fed into the random forest and neural network algorithms. Finally, the performance of algorithms was evaluated using a confusion matrix. Our results showed that the random forest had a promising performance in predicting the amount of required medicine a month ahead using 2 years of data, but its accuracy in predicting medicine type was low (31.8%). We then balanced the dataset and could raise the accuracy of medicine type to 35.9%. Despite the promising accuracy of the random forest model, the required time to generate its model is far more than the NN model.

In the future, we intend to focus only on the requirements for one department of a hospital and use other factors, such as the rarity of medicine, for the prediction task.

**Acknowledgements** Not applicable.

**Author Contributions** The first author designed the prediction method, wrote the manuscript, and conducted the evaluation. The second author collected the dataset and helped in improving the initial idea. The last author revised the manuscript and assisted the first author in the evaluation task.

**Data availability** Not applicable.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Kerman University of Medical Sciences, Kerman, Iran (IR.KMU.REC.1401.076).

## Appendix A

### A.1 Accuracy of the generated models

In Table 11, higher accuracies are bolded. As shown in this table, RF has been superior to NN models in most cases. Sometimes the accuracy of the models is the same up to two decimal places, but considering more decimal places, their accuracy will be slightly different.

### A.2 Required time to generate the models

In Table 12, we presented the required time to build each model. It provides a better understanding of generating the models, and also when the accuracy of models is similar, it can be used as a selection criteria. The required time to generate NN models is significantly less than the RF models.

### A.3 Precision, recall and F1 of NN and RF models

The results presented in Tables 13 and 14 are compatible with the ones in Tables 4 and 5. In Tables 13 and 14, the recall of the NN model is sometimes equal to 100%. It shows that the probability of not predicting the prescribed medicine is very low.

**Table 13** Precision, recall and F1 of NN and RF before balancing

| Test sets | A month | | | A season | | | |
|---|---|---|---|---|---|---|---|
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | |
| *Precision* | | | | | | | |
| RF | **33.33%** | **35.18%** | **28.95%** | **30.83%** | **33.13%** | **31.58%** | Medicine |
| NN | 8.11% | 28.79% | 21.20% | 9.23% | 22.70% | 28.90% | |
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | |
| *Precision* | | | | | | | |
| RF | **77.17%** | **83.63%** | **81.55%** | **100.00%** | 81.35% | 79.12% | Amounts |
| NN | 77.06% | 10.77% | 81.55% | 77.74% | **81.36%** | **79.31%** | |
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | |
| *Recall* | | | | | | | |
| RF | 97.78% | 76.54% | 76.65% | 90.40% | 73.71% | 84.39% | Medicine |
| NN | **100.00%** | **100.00%** | **100.00%** | **100.00%** | **100.00%** | **100.00%** | |
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | |
| *Recall* | | | | | | | |
| RF | 99.85% | 99.73% | 100.00% | 99.91% | 99.97% | 99.10% | Amounts |
| NN | **100.00%** | **100.00%** | **100.00%** | **100.00%** | **100.00%** | **100.00%** | |
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | |
| *F1* | | | | | | | |
| RF | **39.33%** | 43.78% | **40.70%** | **42.40%** | **38.33%** | 42.93% | Medicine |
| NN | 15.01% | **44.71%** | 34.99% | 16.90% | 36.99% | **44.84%** | |
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | |
| *F1* | | | | | | | |
| RF | **87.05%** | **90.98%** | **89.84%** | **87.48%** | 89.70% | **89.10%** | Amounts |
| NN | 87.04% | 19.44% | 89.84% | 87.48% | **89.72%** | 88.46% | |

The best results are bolded

**Table 14** Precision, recall and F1 of NN and RF after balancing

| Test sets | A month | | | A season | | | |
|---|---|---|---|---|---|---|---|
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | |
| *Precision* | | | | | | | |
| RF | **100.00%** | **100.00%** | **100.00%** | **99.15%** | **99.06%** | **97.34%** | Medicine |
| NN | 8.09% | 28.76% | 34.03% | 9.22% | 22.68% | 28.90% | |
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | |
| *Precision* | | | | | | | |
| RF | **64.61%** | **81.89%** | **77.09%** | **76.58%** | **78.61%** | **75.14%** | Amounts |
| NN | 64.16% | 81.82% | 77.09% | 71.90% | 78.57% | 75.10% | |
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | |
| *Recall* | | | | | | | |
| RF | 100.00% | 78.46% | 97.80% | 99.75% | 71.44% | 89.92% | Medicine |
| NN | **100.00%** | **100.00%** | **100.00%** | **100.00%** | **100.00%** | **100.00%** | |
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | |
| *Recall* | | | | | | | |
| RF | 98.81% | 99.86% | 100.00% | 95.77% | 99.94% | 99.64% | Amounts |
| NN | **100.00%** | **100.00%** | **100.00%** | **100.00%** | **100.00%** | **100.00%** | |
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | |
| *F1* | | | | | | | |
| RF | **100.00%** | **45.79%** | **64.57%** | **99.45%** | **53.64%** | **67.98%** | Medicine |
| NN | 14.97% | 44.67% | 50.78% | 16.88% | 36.98% | 44.84% | |
| Training sets | 1 year | 2 years | 3 years | 1 year | 2 years | 3 years | |
| *F1* | | | | | | | |
| RF | 78.13% | 89.99% | 87.06% | **85.11%** | **88.00%** | **87.38%** | Amounts |
| NN | **78.16%** | **90.00%** | **87.06%** | 83.65% | 88.00% | 85.78% | |

The best results are bolded

## References


1. Linnér, L., Eriksson, I., Persson, M., Wettermark, B.: Forecasting drug utilization and expenditure: ten years of experience in Stockholm. BMC Health Serv. Res. **20**(1), 1–11 (2020)
2. Godman, B., Bucsics, A., Vella Bonanno, P., Oortwijn, W., Rothe, C.C., Ferrario, A., Bosselli, S., Hill, A., Martin, A.P., Simoens, S., et al.: Barriers for access to new medicines: searching for the balance between rising costs and limited budgets. Front. Public Health **6**, 328 (2018)
3. Godman, B., Malmström, R.E., Diogene, E., Gray, A., Jayathissa, S., Timoney, A., Acurcio, F., Alkan, A., Brzezinska, A., Bucsics, A., et al.: Are new models needed to optimize the utilization of new medicines to sustain healthcare systems? Expert. Rev. Clin. Pharmacol. **8**(1), 77–94 (2015)
4. Gustafsson, L.L., Wettermark, B., Kalin, M., Korkmaz, S., Persson, M.E., Almkvist, H., Hjemdahl, P., Julander, M., Kristianson, K., Ringertz, B., et al.: A model for structured introduction of new drugs the aim is to offer all patients appropriate treatment. Lakartidningen **105**(42), 2917–2922 (2008)
5. D'Sa, M.M., Nakagawa, R.S., Hill, D.S., Tan, J.K.: Exponential smoothing method for forecasting drug expenditures. Am. J. Health Syst. Pharm. **51**(20), 2581–2588 (1994)
6. Martei, Y.M., Chiyapo, S., Grover, S., Hanna, C., Dryden-Peterson, S., Pusoentsi, M., Shulman, L.N., Tapela, N.: Methodology to forecast volume and cost of cancer drugs in low-and middle-income countries. J. Glob. Oncol. **4**, 1–8 (2018)
7. Remuzat, C., Urbinati, D., Kornfeld, A., Vataire, A.-L., Cetinsoy, L., Aballea, S., Mzoughi, O., Toumi, M.: Pharmaceutical expenditure forecast model to support health policy decision making. J. Mark. Access Health Policy **2**(1), 23740 (2014)
8. Vataire, A.-L., Cetinsoy, L., Aballea, S., Remuzat, C., Urbinati, D., Kornfeld, A., Mzoughi, O., Toumi, M.: Novel methodology for pharmaceutical expenditure forecast. J. Mark. Access Health Policy **2**(1), 24082 (2014)
9. Salmanpour, M.R., Hosseinzadeh, M., Bakhtiari, M., Gholami, A.R., Ghaemi, M.M., Nabizadeh, A.H., Rezaeijo, S.M., Rahmim, A.: Drug amount prediction in Parkinson's disease using hybrid machine learning systems and radiomics features. Soc. Nuclear Med. **63**, 1–4 (2022)
10. Cha, M., Rifai, B., Sarraf, P.: Pharmaceutical forecasting: throwing darts? Nat. Rev. Drug Discov. **12**(10), 737 (2013)
11. Hartke, P.L., Vermeulen, L.C., Hoffman, J.M., Shah, N.D., Doloresco, F., Suda, K.J., Li, E.C., Matusiak, L.M., Hunkler, R.J., Schumock, G.T.: Accuracy of annual prescription drug expenditure forecasts in AJHP. Am. J. Health Syst. Pharm. **72**(19), 1642–1648 (2015)

12. Espin, J., Schlander, M., Godman, B., Anderson, P., Mestre-Ferrandiz, J., Borget, I., Hutchings, A., Flostrand, S., Parnaby, A., Jommi, C.: Projecting pharmaceutical expenditure in eu5 to 2021: adjusting for the impact of discounts and rebates. Appl. Health Econ. Health Policy **16**(6), 803–817 (2018)
13. Keehan, S.P., Stone, D.A., Poisal, J.A., Cuckler, G.A., Sisko, A.M., Smith, S.D., Madison, A.J., Wolfe, C.J., Lizonitz, J.M.: The prediction of essential medicines demand: a machine learning approach using consumption data in Rwanda. Health Aff. **36**(3), 553–563 (2017)
14. Schey, C., Milanova, T., Hutchings, A.: Estimating the budget impact of orphan medicines in Europe: 2010–2020. Orphanet J. Rare Dis. **6**(1), 1–10 (2011)
15. Slade, E.P., Simoni-Wastila, L.: Forecasting medicaid expenditures for antipsychotic medications. Psychiatr. Serv. **66**(7), 713–718 (2015)
16. Rafsanjani, H.N., Nabizadeh, A.H.: Towards human-centered artificial intelligence (AI) in architecture, engineering, and construction (AEC) industry. Comput. Hum. Behav. Rep. **15**, 100319 (2023)
17. Urbinati, D., Rémuzat, C., Kornfeld, Å., Vataire, A.-L., Cetinsoy, L., Aballéa, S., Mzoughi, O., Toumi, M.: Eu pharmaceutical expenditure forecast. J. Mark. Access Health Policy **2**(1), 23738 (2014)
18. Mahajan, K.N., Kumar, A.: Business intelligent smart sales prediction analysis for pharmaceutical distribution and proposed generic model: analyze benefit of institutional level public outreach campaign to business. Int. J. Innov. Res. Technol. **5**(12), 1039–1044 (2019)
19. Chen, J.: Autoregressive integrated moving average (ARIMA). Advanced Technical Analysis concept (2019)
20. Abiodun, O.I., Jantan, A., Omolara, A.E., Dada, K.V., Nachaat AbdElatif, M., Arshad, H.: State-of-the-art in artificial neural network applications: a survey. Heliyon **4**(11), 00938 (2018)
21. Rigatos, G.: Advanced Models of Neural Networks: Nonlinear Dynamics and Stochasticity in Biological Neurons. Springer (2013). https://doi.org/10.1007/978-3-662-43764-3
22. Campos Souza, P.V.: Fuzzy neural networks and neuro-fuzzy networks: a review the main techniques and applications used in the literature. Appl. Soft Comput. **92**, 106275 (2020)
23. Chang, O., Naranjo, I., Guerron, C., Criollo, D., Guerron, J., Mosquera, G.: A deep learning algorithm to forecast sales of pharmaceutical products. no. August (2017)
24. Miikkulainen, R., Liang, J., Meyerson, E., Rawal, A., Fink, D., Francon, O., Raju, B., Shahrzad, H., Navruzyan, A., Duffy, N., Hodjat, B.: Evolving deep neural networks. In: Artificial Intelligence in the Age of Neural Networks and Brain Computing, pp. 293–312. Elsevier (2019). https://doi.org/10.1016/B978-0-12-815480-9.00015-3
25. Kwak, Y., Yun, W.J., Jung, S., Kim, J.: Quantum neural networks: concepts, applications, and challenges. In: 2021 Twelfth International Conference on Ubiquitous and Future Networks (ICUFN), pp. 413–416 (2021). IEEE
26. Kumar, S.A., Kumar, A., Dutt, V., Agrawal, R.: Multi model implementation on general medicine prediction with quantum neural networks. In: 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), pp. 1391–1395. IEEE (2021)
27. Meyer, D., Wien, F.: Support vector machines. The Interface to libsvm in package e1071 28(20), 597 (2015)
28. Yang, F.-J.: An implementation of Naive Bayes classifier. In: 2018 International Conference on Computational Science and Computational Intelligence (CSCI), pp. 301–306. IEEE (2018)
29. Biau, G., Scornet, E.: A random forest guided tour. TEST **25**, 197–227 (2016)
30. Alves, C.M.F.: Demand forecasting in a multi-specialty hospital setting: a comparative study of machine learning and classical statistical methods. Mestrado Integrado em Engenharia Informatica e Computação **Faculdade De Engenharia Da Universidade Do Porto** (2018)
31. Sherstinsky, A.: Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D **404**, 132306 (2020)
32. Kalekar, P.S., et al.: Time series forecasting using holt-winters exponential smoothing. Kanwal Rekhi Sch. Inf. Technol. **4329008**(13), 1–13 (2004)
33. Bakırarar, B., Yüksel, C., Yavuz, Y.: Prediction of new prescription requirements for diabetes patients using big data technologies. J. Health Res. **36**(2), 334–344 (2021)
34. Taud, H., Mas, J.: Multilayer perceptron (MLP). Geomatic approaches for modeling land change scenarios, pp. 451–455 (2018)
35. Gasso, G.: Logistic regression. INSA Rouen-ASI Departement Laboratory: Saint-Etienne-du-Rouvray, France, pp. 1–30 (2019)
36. Bhargava, N., Sharma, G., Bhargava, R., Mathuria, M.: Decision tree analysis on j48 algorithm for data mining. Proc. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **3**(6), 15 (2013)
37. Mbonyinshuti, F., Nkurunziza, J., Niyobuhungiro, J., Kayitare, E.: The prediction of essential medicines demand: a machine learning approach using consumption data in Rwanda. Processes **10**(1), 26 (2021)
38. Mbonyinshuti, F., Nkurunziza, J., Niyobuhungiro, J., Kayitare, E.: Application of random forest model to predict the demand of essential medicines for non-communicable diseases management in public health facilities. Pan Afr Med. J. **42**, 15 (2022)
39. Rashid, M., Goyal, V., Parah, S.A., Singh, H.: Drug prediction in healthcare using big data and machine learning. In: Research Anthology on Machine Learning Techniques, Methods, and Applications, pp. 1062–1071. IGI Global (2022)
40. Guseo, R., Dalla Valle, A., Furlan, C., Guidolin, M., Mortarino, C.: Pre-launch forecasting of a pharmaceutical drug. Int. J. Pharm. Halthc. Mark. **11**(4), 412–438 (2017)
41. Nabizadeh, A.H., Gonçalves, D., Gama, S., Jorge, J.: Early prediction of students's final grades in a gamified course. IEEE Trans. Learn. Technol. **15**(3), 311–325 (2022)
42. Demisse, G.B., Tadesse, T., Bayissa, Y.: Data mining attribute selection approach for drought modeling: a case study for greater horn of Africa. arXiv preprint arXiv:1708.05072 (2017)
43. Ying, X.: An overview of overfitting and its solutions. In: Journal of Physics: Conference Series, vol. 1168, p. 022022. IOP Publishing (2019)
44. Deng, K.: Omega: On-line Memory-based General Purpose System Classifier. Carnegie Mellon University (1999)
45. Salmanpour, M.R., Hosseinzadeh, M., Bakhtiari, M., Ghaemi, M.M., Rezaeijo, S.M., Nabizadeh, A.H., Rahmim, A.: Cognitive outcome prediction in Parkinson's disease using hybrid machine learning systems and radiomics features. Soc. Nuclear Med. (2022)
46. Javanmardi, A., Hosseinzadeh, M., Hajianfar, G., Nabizadeh, A.H., Rezaeijo, S.M., Rahmim, A., Salmanpour, M.: Multi-modality fusion coupled with deep learning for improved outcome prediction in head and neck cancer. In: Medical Imaging 2022: Image Processing, vol. 12032, pp. 664–668. SPIE (2022)
47. Miao, J., Niu, L.: A survey on feature selection. Procedia Comput. Sci. **91**, 919–926 (2016)
48. Kursa, M.B., Rudnicki, W.R.: Feature selection with the Boruta package. J. Stat. Softw. **36**, 1–13 (2010)
49. Genuer, R., Poggi, J.-M., Genuer, R., Poggi, J.-M.: Random Forests. Springer, New York (2020)
50. Nabizadeh, A.: Artificial intelligence in healthcare. J. Health Biomed. Inform. **9**(3), 193–195 (2022)
51. Qi, Y.: Random Forest for Bioinformatics. Ensemble machine learning: Methods and applications, pp. 307–323. Springer, Berlin (2012)

52. Subbotin, S.: A random forest model building using a priori information for diagnosis. In: CMIS, pp. 962–973 (2019)
53. Salmanpour, M.R., Hosseinzadeh, M., Akbari, A., Borazjani, K., Mojallal, K., Askari, D., Hajianfar, G., Rezaeijo, S.M., Ghaemi, M.M., Nabizadeh, A.H., et al.: Prediction of TNM stage in head and neck cancer using hybrid machine learning systems and radiomics features. In: Medical Imaging 2022: Computer-Aided Diagnosis, vol. 12033, pp. 648–653. SPIE (2022)
54. Mijwel, M.M.: Artificial neural networks advantages and disadvantages. Retrieved from LinkedIn, https://www.linkedin.com/pulse/artificial-neural-networks-advantages-disadvantages-maad-m-mijwel, pp. 21 (2018)
55. Raleiras, M., Nabizadeh, A.H., Costa, F.A.: Automatic learning styles prediction: a survey of the state-of-the-art (2006–2021). J. Comput. Educ. **9**(4), 587–679 (2022)
56. Nabizadeh, A.H.: A long term goal recommender approach for learning environments. Doutoramento em Informática **Faculdade De ciencias Da Universidade Do Porto** (2018)
57. Nabizadeh, A.H., Jorge, J., Gama, S., Goncalves, D.: How do students behave in a gamified course? A ten-year study. IEEE Access **9**, 81008–81031 (2021)
58. Sadoughi, F., Ghaderzadeh, M.: A hybrid particle swarm and neural network approach for detection of prostate cancer from benign hyperplasia of prostate. In: e-Health–For Continuity of Care, pp. 481–485. IOS Press (2014)
59. Fernández, A., Garcia, S., Herrera, F., Chawla, N.V.: Smote for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. J. Artif. Intell. Res. **61**, 863–905 (2018)
60. Iqbal, M.J., Geer, M.I., Dar, P.A.: Evaluation of medicines forecasting and quantification practices in various public sector hospitals using indicator based assessment tool. J. Appl. Pharm. Sci. **7**(12), 72–76 (2017)