A clinical support system analysing drug prescriptions based on AI, in order to limit drug iatrogeny in hospitals (August 2022)

First A. Author, *Fellow, IEEE*, Second B. Author, and Third C. Author, Jr., *Member, IEEE*

***Abstract*— Context: The health product circuit corresponds to the chain of steps that a medicine goes through in hospital, from prescription to administration. The safety and regulation of all the stages of this circuit are major issues to ensure the safety and protect the well-being of hospitalized patients. Indeed, each stage involves risks that can lead to iatrogenic medication, which affects between 6 and 10% of hospitalised patients. Our study is done in collaboration with the Centre Hospitalier Universitaire de Lille in France.**

Objective: This paper presents an automatic system for analysing prescriptions using AI and ML, with the aim of ensuring patient safety by limiting the risk of prescription errors.

Materials and Methods: Our pharmaceutical prescription analysis system uses the MIMIC-III database containing information on each prescription (drug molecule, dosage, date, etc.), on the patient (age, gender, etc.) and on the patient's biological tests (renal function, blood platelet count, etc.). With our system, the validity of the prescription as well as the compatibility with the patient's condition and tests is done by a clustering method and by Machine Learning methods. Our system allows not only a simple analysis of each individual prescription, but also an inter-drug analysis between all medications taken by a patient during the same hospital stay.

Results: We tested our methods on the positive data, i.e. conforming, from MIMIC-III as well as on the negative data, i.e. problematic, that we created. The clustering methods give an accuracy of 0.4391 and an F1 Score of 0.3544. Machine learning methods give excellent results, with an accuracy of 0.9942 and an F1 Score of 0.9942.

Discussion : On the one hand, the mixed results of clustering methods can be explained by several leads, in particular the Euclidean calculation of distances. It would be interesting to compute distances with different coefficients for each column, in order to give more weight to the most important columns. On the other hand, the results of supervised learning methods (XGBoost and Random Forest) are excellent. The algorithms highlight that three parameters are much more important than the others in the prescription classification process: the dose of medication prescribed to the patient, the age of the patient, and the patient's creatinine level in the blood, which indicates a potential renal insufficiency. Knowing this is valuable and would allow us to refine the coefficients in the recalculation of distances.

Conclusion : Our study in collaboration with the University Hospital of Lille allowed us to explore different algorithmic and statistical methods to analyze pharmaceutical prescriptions, with the aim of limiting drug iatrogeny. For this work, we exploited the MIMIC-III database and proposed three approaches. The first one, clustering methods, proved to be moderately efficient. We have several avenues of improvement to improve our clustering methods, including calculating distances between prescriptions differently, and testing the algorithms on other negative data that will be real. The second approach, Machine Learning, gave excellent results and allowed us to discern the most important features in the prescription classification process. The third approach consists in identifying potential drug interactions between two drugs prescribed to the same patient during the same hospital stay. To do this, we check whether or not these two drugs have been prescribed together before in our database. Concerning the future perspectives, it would be interesting to explore the tracks explained before to improve the clustering methods. It would also be interesting to test our methods of analysis of potential drug interactions.

***Index Terms*— Artificial Intelligence ; Decision support system; Drug interactions; Medical prescription review; Pharmacy;**

I. Introduction

Iatrogeny refers to all disorders, undesirable effects or illnesses caused by a medical intervention, treatment or medication and represents a major public health issue, both in terms of health and economics. Medication-related iatrogeny, induced by the administration of one or more drugs, concerns 6 to 10% of hospitalised patients and is responsible for 3.5% of deaths (DGOS report, 2004). At the global level, in a document entitled "To err in human", the IOM (1999) estimated iatrogeny as equivalent to a Boeing 747 crashing every two days. In a French study (2009), it was shown that 33% of serious adverse drug reactions were drug-related and that 51.2% could have been avoided (DRESS, 2009). Apart from the health impact, this iatrogeny has an important economic impact on health institutions and therefore on society. In its 2006 report, the Council of Europe estimated the annual cost of adverse drug reactions at 636 million euros, of which 38% were avoidable.

In this context, making therapeutic management safe is a priority for health care institutions in order to reduce iatrogeny by acting simultaneously on the various stages of the health product circuit: prescription, transcription, delivery and administration. The safety of each stage is a national objective which aims to ensure the 5B rule: the right patient receives the right drug or device, in the right dose, at the right time, by the right route.

In hospitals, the pharmacy analyses prescriptions before dispensing medicines to patients in order to detect possible errors and ensure patient safety. There are many possible errors in a prescription: unusual dosage for a specific active ingredient, inappropriate route of administration, overdose in relation to the patient's age, toxic interaction between two drugs on the same prescription, etc. The analysis of prescriptions is therefore complex and must absolutely take into account the following three criteria: the prescription itself (dosage, route of administration, frequency of administration), compatibility with the patient's biological characteristics and analyses (age, weight, renal function, etc.) and finally possible drug interactions.

Currently, it is most often the pharmacists themselves who proofread the prescriptions. However, this task is tedious and time-consuming. In most hospitals, such as the University Hospital of Lille, the pharmacy is not able to review all prescriptions: only certain

This paragraph of the first footnote will contain the date on which you submitted your paper for review. It will also contain support information, including sponsor and financial support acknowledgment. For example, "This work was supported in part by the U.S. Department of Commerce under Grant BS123456."

The next few paragraphs should contain the authors' current affiliations, including current address and e-mail. For example, F. A. Author is with the National Institute of Standards and Technology, Boulder, CO 80305 USA (e-mail: author@boulder.nist.gov).

S. B. Author, Jr., was with Rice University, Houston, TX 77005 USA. He is now with the Department of Physics, Colorado State University, Fort Collins, CO 80523 USA (e-mail: author@lamar.colostate.edu).

T. C. Author is with the Electrical Engineering Department, University of Colorado, Boulder, CO 80309 USA, on leave from the National Research Institute for Metals, Tsukuba, Japan (e-mail: author@nrim.go.jp).

departments and patients, notably incoming patients, benefit from this pharmaceutical verification.

However, the computerisation of patient records has paved the way for the automation of the review of pharmaceutical prescriptions by AI. Several IT tools and expert systems already exist to assist pharmacists in reviewing prescriptions (detailed later in the state of the art). However, these tend to generate many unnecessary alerts and suggest problematic alternative treatments. In addition, most of these systems do not take drug interactions into account well.

In this context, we have developed an AI module to analyse the compliance of pharmaceutical prescriptions, in cooperation with the Lille University Hospital. This system exploits a database containing the information of each prescription, as well as the characteristics and biological analyses of the patient. The validity of the prescription, as well as its compatibility with the patient's condition and analyses, is checked using a clustering method. Drug interactions are checked using word embedding.

II. State of the Art

This state of the art addresses the whole of the health products circuit, it is difficult to make an exhaustive synthesis of the state of the art, as the field is so vast. In this article, we will try to summarise the key points of this circuit.

A. Storage

The aim of automation is to achieve the tightest possible control over the use of medicines in the healthcare system. A number of widely cited studies have found that hospital patients are systematically at risk of wrong dose, wrong time or wrong drug errors [3], [4], [5]. These studies demonstrate the emerging role of clinical pharmacy as an agent of drug use control and underscore its claim to cultural authority as the guardian of medication safety for the health care system. In order to free up pharmacists' time for clinical tasks while reducing the risk of medication errors, from the 1960s onwards, a new design of drug distribution systems was introduced. Although widely used in North America, automated drug distribution systems are not yet widely deployed in Europe [6]. Unit-dose drug distribution anticipates and parallels modern management techniques derived from Japan, such as just-in-time management or total quality management [7], which seek to reduce costs and improve quality by reducing stocks and eliminating buffer stocks of materials.

Drug distribution systems (DDS) are computerised drug storage devices or cabinets that provide computer-controlled storage, delivery and tracking of medicines. These systems have been identified as a potential mechanism for improving efficiency and patient safety [8]. The efficiency of the drug dispensing process is improved by automating the tasks of pharmaceutical workers that require repetitive movements, high concentration and reliable record keeping. Dispensing time is reduced, as stock management is dictated by pre-set minimum and maximum levels. A study to evaluate the effect of implementing an automated dispensing system in the pharmacy of Valenciennes hospital showed a significant decrease in dispensing errors [9]. However, in some cases, errors may even increase with the use of SDM technology. The United States Pharmacopeia-Institute for Safe Medication Practices Medication Error Reporting Program (USP- ISMP MERP) has also reported medication errors associated with the use of MDS [10]. The main errors were due to the computer interface, poor communication with other systems and lack of sufficient decision support. Human error also played a role in the errors through inexperience, inadequate knowledge, interruptions and typing errors. As with any new technology, automated delivery systems can compromise patient safety if they are not managed appropriately. However, the human agents involved in the medication process are able to detect, correct and learn from errors. They have a key role to play in addressing safety issues by ensuring that MDS are used in a way that improves the safe use of medicines, and by developing patient safety initiatives such as medication reconciliation. The safe use of MDS therefore depends on human-machine systems and human-human cooperation between health professions to plan and implement a medication safety system that incorporates automation technology.

B. Medical devices

In its two successive versions (2004 and 2009), the national ENEIS study highlighted serious avoidable adverse events linked to their use [11]. Few studies have attempted to evaluate the iatrogenic costs associated with medical devices, as the field is so vast. Nevertheless, we can cite the APHP [12] which estimates a minimum amount of €1.51/day, i.e. extrapolated at national level to €43 million/year, excluding unreported events, extra-hospital costs and long-term consequences for the patient. More generally, the authorities point to "uncontrolled expenditure" [13].

To date, there are more than twenty thousand generic device groups in nomenclatures such as GMDN (global medical device nomenclature) or CND (the European nomenclature for medical devices). The CND nomenclature adopted by Europe, originally a specific nomenclature for Italy, was adopted as part of the implementation of the new European regulation. Italy is in fact about ten years ahead in the management of medical devices and is currently leading the European work in this area. Thus, the technique of identification by radio frequency RFID has been developed there. The object to be traced is tagged with an electronic chip associated with an antenna. Using a suitable reader, it is possible to 'stimulate' the chip in order to read the information it contains from a distance.

C. Analysis of the prescriptions

Pharmaceutical analysis of drug prescriptions is an essential step in the drug circuit, in securing drug management (indicator of the Contract for Improved Quality and Efficiency of Care) and conditions the act of dispensing a drug (Decree n°2019-489 of 21 May 2019). This clinical pharmacy activity [14] is more or less developed in French health care institutions, in particular because it is a relatively time-consuming activity in relation to the rates of pharmaceutical interventions observed [15], [16], [17] and a high requirement in terms of training and updating of knowledge. With the computerisation of patient records and drug prescriptions, and the steps taken to standardise data so that they can be exchanged between different computer systems, pharmacists have an opportunity to subcontract part of this prescription analysis activity to artificial intelligence.

Several studies have already been published on the development of alert rules [18] and computerised clinical decision support (CDS) alert systems. However, these CDS alert systems tend to send many unnecessary alerts, which leads to alert fatigue in healthcare [19]. This alert fatigue can be counterproductive and dangerous, as pharmacists end up no longer systematically checking prescriptions deemed problematic by the CDS and therefore miss out on justified alerts.

Other studies are interested in patient and drug profiling to refine the analysis of prescriptions and target the patients and drugs most at risk. The article by C. Jarre et al [20] shows that the most relevant parameters are biological criteria and the use of certain drug families.

In recent years, Machine Learning and AI have started to be used to develop decision support systems to identify prescriptions with a high risk of leading to drug iatrogeny. One example is the paper by J. Corny et al. whose decision support system uses a hybrid approach (machine learning and rule-based expert system). The results of this hybrid system are very promising: according to the article, the areas under the receiving-operating characteristic and precision-recall curves of their system reached 0.81 and 0.75 respectively, whereas these metrics were equal to 0.65 and 0.56 respectively for the CDS system, and were equal to 0.68 and 0.56 respectively for the multi-criteria query techniques.

Some hospitals have embarked on the development of local solutions, such as the Charité sur Loire Hospital [21] (Lagrange F et al., 2017), which has created an alert system based on a set of rules and checks targeted at the elderly, to limit drug iatrogeny among seniors. This system has changed the prescription rates of different categories of drugs in the elderly. An initial assessment of the effectiveness of the system was made by observing the number of falls among the elderly, which fell from 57 out of 185 patients before the alert system was introduced to 34 falls out of 184 patients using their alert system. These results should be taken with caution, as falls in the elderly are multi-causal.

Today, various companies have launched pharmaceutical decision support systems to assist the clinical pharmacist in this activity. One example is the French company Keenturtle, which has developed AI-based prescription analysis software called ”PharmaClass”. This software takes into account the prescription itself, the drug encyclopaedia listing the information relating to each drug, the patient’s data (age, sex, allergies, illnesses etc.) and the patient’s biological analyses. The PharmaClass software is currently used by several hospitals in France, including the Lille University Hospital. Another example is PharmIA, a Quinten Group company that has developed an intelligent digital platform for hospital pharmacists to limit drug iatrogeny. Their system uses expert rules, as well as drug regulations and patient data (clinical, biological and history data). Other software exists and is used in the hospital environment, such as VIDAL Sentinel or Bimedoc Expert.

III. MATERIALS AND METHODS

A. Initial data MIMIC-III

For this article, we used the MIMIC-III data (’Medical Information Mart for Intensive Care’) which is a database accessible to all researchers worldwide after completing an appropriate online training and signing a data use agreement. The open nature of this database makes it perfect for making studies reproducible and for conducting academic or industrial research projects.

The MIMIC-III data include deidentified clinical data on 38,597 adult patients admitted between 2001 and 2012 to the intensive care unit at Beth Israel Deaconess Medical Center in Boston, Massachusetts. The data also include 7870 neonates admitted between 2001 and 2008.

B. Data formatting

MIMIC-III data are grouped into 26 different tables, grouping different types of medical data. Our work focuses on the PRESCRIPTIONS table.

First, we merge the PATIENTS table with the PRESCRIPTIONS table. Then, for each prescription row in this new table, we add the laboratory measurements from the LABEVENTS table that share the same hospital stay identifier as the prescription.

Thus, we obtain a single dataframe where each line corresponds to a prescription, enriched with the data of the patient concerned (gender, age, etc.) as well as the biological analyses of this one.

We deleted all prescription lines that were missing the drug or dose. Regarding the biological analyses, we kept those that were both relevant, and measured for most patients, to avoid too much NaN. The biological assays kept for our dataframe are: *Anion Gap, Bicarbonate, Calcium (Total), Chloride, Creatinine, Glucose, Hematocrit, Hemoglobin, Mean Corpuscular Hemoglobin, Mean Corpuscular Hemoglobin Concentration, Mean Corpuscular Volume, Magnesium, Phosphate, Platelet Count, Potassium, Red Cell Distribution Width, Red Blood Cells, Sodium, Urea Nitrogen and White Blood Cells*.

We divided the resulting DataFrame into several DataFrames, separating them by drug as well as by route of administration and prescription unit. Not distinguishing the different routes of administration or the prescription units would disturb the clustering algorithms because comparisons between routes of administration for the same active ingredient are not always possible. Indeed, how can we compare the effects of the same drug taken orally or by the skin? How to compare the therapeutic effect of a dose in grams and a dose in inhaled puffs?

C. Additional data created

The prescriptions in the MIMIC-III database can be considered as valid and normal prescriptions, which would all be validated by a pharmacist. We therefore only have validated prescriptions, the ”positives”. We must therefore create ”negatives”, i.e. problematic prescriptions that would be refused by a pharmacist, if we want to be able to use supervised learning methods as well.

To create these ”negative” or problematic prescription rows, we start with prescription rows from the MIMIC database and then randomly change the numerical values. To ensure that the data created are indeed ”negatives”, the numerical values in each changed column must not belong to a confidence interval. The confidence interval is the mean of the column $\pm$ the standard deviation of the column.

The negative prescriptions created are individually problematic prescriptions, without considering potential drug interactions. The negative prescriptions created should therefore only be used to test methods that analyze each prescription individually.

D. Preprocessing

We used classical preprocessing methods, such as deleting columns not needed for our study and unifying the units of measurement. For each column, we replace missing values with a random variable that follows a Gaussian distribution with mean equal to the column mean and standard deviation of the column. We also perform an encoding for the categorical variables. Finally, we proceed to a normalization step.

E. Methods of analysis of each prescription individually

We have developed several methods to study each prescription individually: clustering methods and classical Machine Learning methods.

1) Clustering methods: To create the clusters, only MIMIC data is used, i.e. valid and compliant prescriptions. For each training dataframe corresponding to a drug (D), a prescription unit (PU) and a specific route of administration (RA), we cluster the data, with a k-means clustering model and with a hierarchical clustering model. Each point of the cluster corresponds to a prescription corresponding to this D, PU and RA.

The k-means clustering algorithm takes as input the data to be clustered and a number of clusters to be formed. When given Nclusters=k as input to the Kmeans algorithm, the algorithm creates k random centroids. Then for each point of the input data, the algorithm looks at which centroid it is closest to, and associates it with this centroid. Then, the coordinates of each centroid are modified: each centroid becomes the average point of all the points associated to this centroid. We repeat the previous steps until the centroids do not change anymore.

The hierarchical clustering model has a major difference with the k-means model: it does not need to take as an argument the number of clusters to be formed. Indeed, it calculates itself the number of clusters. It takes as input a distance from which two clusters should not be merged. Hierarchical Clustering starts by considering each point as a cluster. Then, the algorithm finds the two closest clusters to each other, and merges them into one cluster. The algorithm repeats this step as long as some clusters are less distant from each other than the threshold distance entered as an argument.

To test if a new prescription (corresponding D, PU and RA) is compliant, we look at which cluster it should belong to. Then we compare the euclidean distance between the new prescription and the center of the cluster with the diameter (i.e. the maximum distance between a cluster point and its center). If the new prescription to be tested is further from the cluster center than all the other points belonging to the cluster, then the prescription is considered problematic and non-conforming.

Here is the alert algorithm that analyzes a new prescription to be studied by also taking as input the parameters of the clustering model used and trained, and returns a boolean: True indicates an alert, False indicates that the prescription is normal.

Algorithm 1 Warning system for clustering method

```
function WARNING(line: Series, model_params: dict)
    mean ← model_params['mean']
    std ← model_params['std']
    point ← normalize_line(line.values, mean, std)
    distance_to_clusters ← [euclidean(center, point) for center in centers] ▷ We calculate the distance between the point and each cluster
    index_center ← np.argmax(distance_to_clusters) ▷ We select the index of the smallest distance calculated in the previous line
    center ← centers[index_center] ▷ 'center' is the cluster center closest to the point
    distance ← euclidean(center, point)
    maximum_distance ← model_params['max_dist'][index_center]
    return distance < maximum_distance
```

Clustering methods analyze the validity of a prescription individually, without taking into account possible drug interactions with other prescriptions of the patient for the same hospital stay.

2) Machine Learning supervised methods: For these supervised Machine Learning methods, the training data are a homogeneous mix of conforming prescriptions, i.e. positives from MIMIC-III, as well as negative data created. We have chosen to use two classification methods to discriminate the prescriptions as positive, i.e. conforming, or negative, i.e. problematic: the XGBoost method and the Random Forest method.

Like Clustering methods, Machine Leaning methods analyze the validity of a prescription individually, without taking into account possible drug interactions with other prescriptions of the patient for the same hospital stay.

F. Methods of drug interaction analysis

In addition to studying each prescription individually, it is interesting to study possible drug interactions with other drugs prescribed to the same patient during the same hospital stay.

To analyze potentially toxic drug interactions, we start by listing in a list of lists (LL) the drugs that are compatible, i.e. the drugs that have been prescribed to the same patient, in the same hospital stay, during the same period.

To test whether a new prescription could conflict with other drugs that the patient is already taking, we start by listing all the drugs prescribed to this patient during the same hospital stay as the prescription to be tested. Then, we check if the drug to be tested has already been prescribed with these drugs in the LL. If the answer is no, then an alert is returned to warn of the potential risk of drug interaction.

We could also ask doctors for lists of drugs that are compatible with each other, as well as drugs that interact toxically with each other.

IV. Results

A. Results of Clustering Methods

1) Numerical results of clustering methods: To study the results of the clustering methods, we focused on the clustering of the drug ACET325, with the prescription unit mg and with the administration route 'PO'. We chose this drug because the MIMIC database contains more than 30000 prescriptions of this drug, which allows a serious study. ACET325 is an analgesic and antipyretic, often prescribed to relieve pain caused by various conditions such as headaches and osteoarthritis, and to reduce fever caused by infections.

We first tested our clustering method with the Kmeans algorithm on the ACET325 prescriptions. We varied the number of clusters parameter and obtained the best results for 5 clusters. We obtained the following confusion matrix:

$$Confusion\ Matrix\ Kmeans = \begin{pmatrix} 2784 & 6362 \\ 3930 & 5215 \end{pmatrix} \quad (1)$$

We then tested our clustering method with the Hierarchical algorithm on the ACET325 prescriptions. This algorithm indicates that it finds more than 2800 clusters among these prescriptions. We obtained the following confusion matrix:

$$Confusion\ Matrix\ Hierarchical = \begin{pmatrix} 2916 & 6360 \\ 3929 & 5216 \end{pmatrix} \quad (2)$$

These two confusion matrices, Kmeans and Hierarchical, allowed us to calculate the Accuracy, Precision, Recall and F1 Score for these two methods:

TABLE I
NUMERICAL RESULTS OF THE CLUSTERING METHODS

Score	Kmeans method	Hierarchical method
Accuracy	0.4373	0.4391
Precision	0.3043	0.3078
Recall	0.4146	0.4174
F1score	0.3510	0.3544

2) Visualization of clusters: Since there are a substantial number of combinations between the models and drugs to be tested for visualization, we decided to do the visualization for the Kmeans model with ACET325. Since each data point (each representing a prescription) has a dimension of 25 (of the 25 columns used), we used Principal Component Analysis for dimension reduction.

We visualized in 2D and 3D the clusters created by the Kmeans algorithm to discriminate ACET325 prescriptions, for 2 clusters and then 5 clusters. The lines are connecting each prescription point to the center of their cluster.

Here is the 3D graph displaying the prescriptions of ACET325 divided into 2 clusters calculated by the Kmeans algorithm:

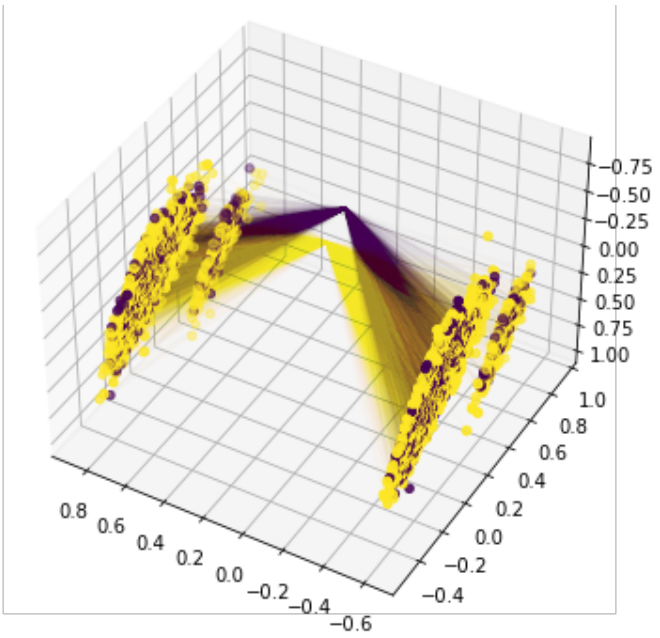

Fig. 1. 3D visualization of clusters of ACET325 prescriptions; number of clusters = 2

Here is the 2D graph displaying the prescriptions of ACET325 divided into 2 clusters calculated by the Kmeans algorithm:

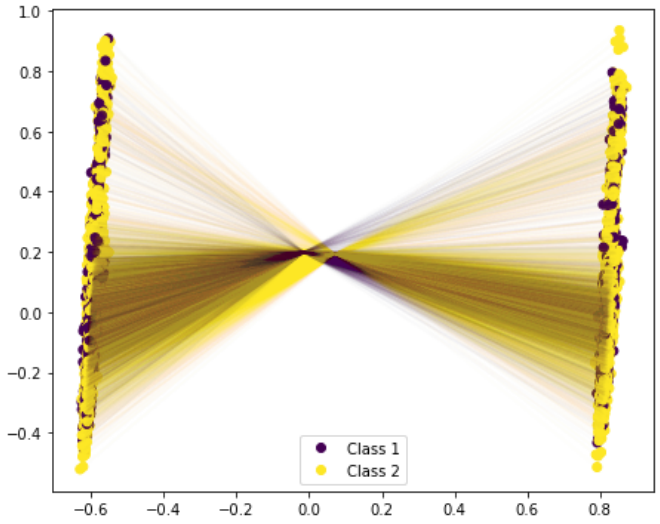

Fig. 2. 2D visualization of clusters of ACET325 prescriptions; number of clusters = 2

Here is the 3D graph displaying the prescriptions of ACET325 divided into 5 clusters calculated by the Kmeans algorithm:

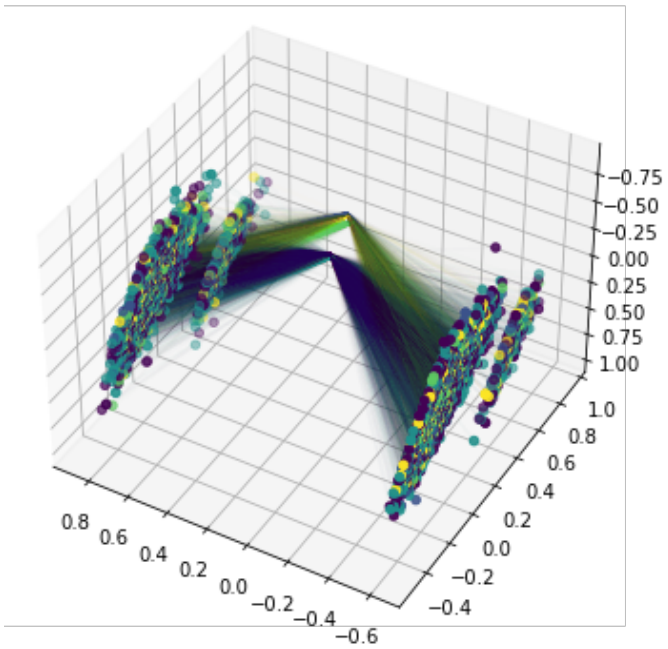

Fig. 3. 3D visualization of clusters of ACET325 prescriptions; number of clusters = 5

Here is the 2D graph displaying the prescriptions of ACET325 divided into 5 clusters calculated by the Kmeans algorithm:

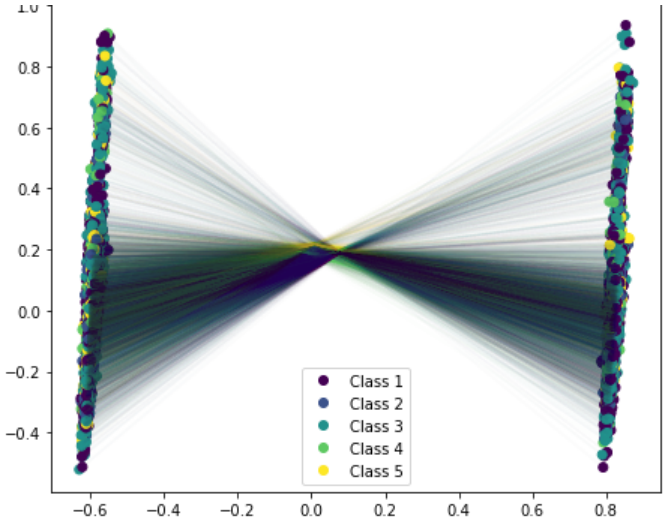

Fig. 4. 2D visualization of clusters of ACET325 prescriptions; number of clusters = 5

B. Results of Machine Learning Methods

In parallel to the clustering methods, we have tested two supervised learning methods: the XGBoost method and the Random Forest method. For these methods, we used for the training as well as for the test, a mix of equal proportions between the positive prescriptions of the MIMIC database and between the negative prescriptions that we created.

1) Results of the XGBoost Method: Our tests with the XGBoost method gave results shown in the following confusion matrix:

$$Confusion\ Matrix\ XGB = \begin{pmatrix} 7096 & 2050 \\ 1681 & 7465 \end{pmatrix} \quad (3)$$

Here is the importance of each column, each feature, in the classification of prescriptions according to the XGBoost algorithm:

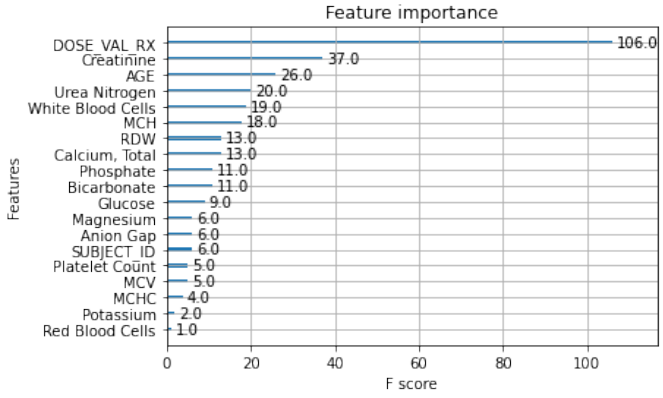

Fig. 5. Importance of features in the classification of prescriptions, according to the XGBoost algorithm

We see that here the most important feature by far is the dose of drug prescribed, followed by the creatinine level and the age of the patient.

2) Results of the Random Forest Method: We then tested the Random Forest method, which yielded the confusion matrix below:

$$Confusion\ Matrix\ RF = \begin{pmatrix} 9060 & 86 \\ 20 & 9126 \end{pmatrix} \quad (4)$$

Here is the importance of each column, each feature, in the classification of prescriptions according to the Random Forest algorithm:

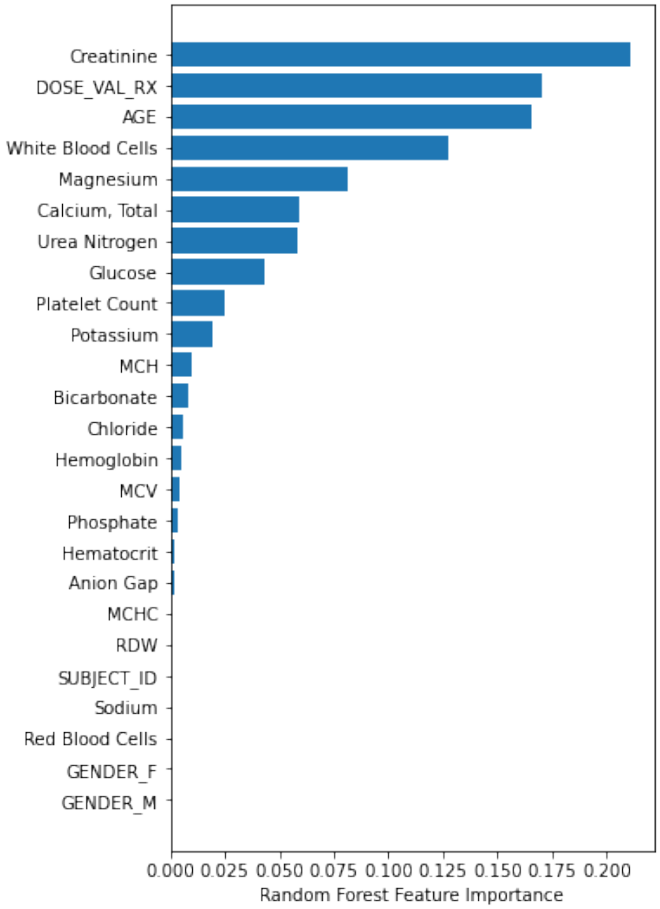

Fig. 6. Importance of features in the classification of prescriptions, according to the Random Forest algorithm

We notice that the three features judged the most important by Random Forest are the same as those of XGBoost, in a different order: creatinine level, then the dose of medication prescribed, then the age of the patient.

The Random Forest method uses several decision trees. Here is one of the decision trees of our Random Forest model trained on our data:

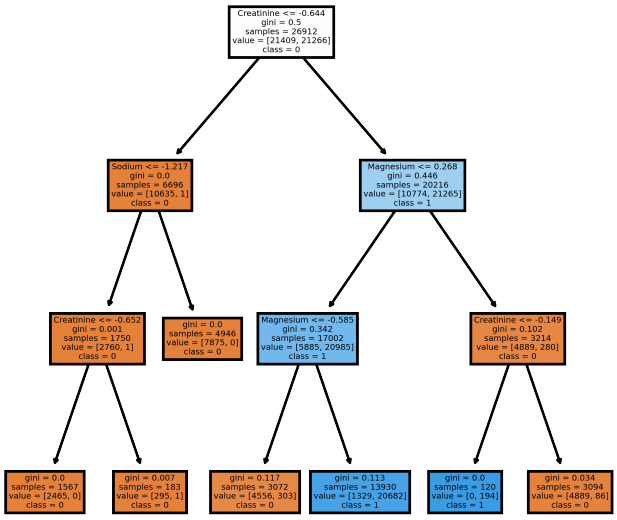

Fig. 7. Importance of features in the classification of prescriptions, according to the Random Forest algorithm

3) Comparison: The two confusion matrices, calculated by the XGBoost and Random Forest methods, are used to calculate the Accuracy, Precision, Recall and F1 Score for these two methods:

TABLE II
RESULTS OF THE MACHINE LEARNING METHODS

Score	XGBoost method	Random Forest method
Accuracy	0.7960	0.9942
Precision	0.7845	0.9906
Recall	0.8162	0.9978
F1score	0.8000	0.9942

C. Results of methods of drug interaction analysis

It is important to remember that we created the negative data by changing for each prescription row randomly the value of each column, making sure that this value does not belong to a confidence interval. Our created negative data can therefore be used to test Clustering and Machine Learning methods, as they are individually negative prescriptions in themselves. However, the negative data we created are not suitable for testing drug interactions between prescriptions.

To test for drug interactions, new negative data could be created by creating several incompatible drug prescriptions for the same patient and hospital stay. To do this, we would need the help of the medical profession, which could give us a list of incompatible drugs for example.

V. DISCUSSION

A. Discussion regarding the results of the Clustering Methods

1) About Numerical Results: We see that the results of the clustering methods are not quite up to our expectations. Indeed, none of the Accuracy, Precision, Recall, F1 Score scores exceed 0.5. Also, we find that for both positive and negative data, the algorithm tends to over-categorize prescriptions as negative. These mixed numerical results can be explained in several ways.

First, calculating distances in a Euclidean way may not be relevant to our problem. Indeed, by using the Euclidean distance, we give equal importance to each column of the data. However, some features could be taken into account more than others, such as the prescribed dose, the creatinine rate or the age of the patient. A future track would be to determine the weight of each feature with the help of the medical profession and the Lille University Hospital, for a more adapted calculation of the distances between the prescriptions.

Second, it is possible that the test would have worked better with real negative data, not created by us for the purposes of our study. Indeed, we created the negative data, i.e. the non-compliant prescriptions, by taking for each column random values outside a confidence interval, i.e. randomly very high or very low values. Each negative prescription therefore has statistically about half of its columns very low, and the other half very high. Thus, the artificially created negative data may not be that far away, in terms of distance calculation, from the positive data provided by MIMIC-III.

2) About the visualisation of Clusters: It is necessary to keep in mind that the prescriptions have 25 columns each. Each point of the cluster has therefore a dimension of 25. Thus, it is normal not to see visually a clear coherence and an assertive geometrical separation of the clusters in the 2D and 3D visualization.

The 2D and 3D visualization remains interesting because it shows in particular that the centers of the various clusters are well separated and not almost confused/overlapped.

B. Discussion regarding the results of the Machine Learning Methods

The results of our supervised learning methods, XGBoost and Random Forest, are excellent, with scores such as the very high F1 Score.

The graphs showing the importance of each feature in the classification process of the two methods clearly show that there are three parameters that matter more: the dose of drug prescribed, the age of the patient, and the creatinine level.

These three features are perfectly logical and relevant to classify the prescriptions.

Indeed, the dose prescribed to the patient is intuitively the central parameter: all drugs have maximum doses that should not be exceeded.

Moreover, the recommended dose varies greatly depending on the age of the patient: very young and very old patients often have to ingest lower maximum amounts than the one recommended for healthy adults.

In addition, many drugs cannot be prescribed if the patient has weak kidneys. Creatinine levels are an indicator of kidney health: high creatinine levels indicate kidney failure.

It is valuable to know which features are most important in the prescription classification process. This could lead in the future to give more weight to e.g. the prescribed dose, the patient's age and his creatinine level in the calculation of the distances between prescriptions.

VI. CONCLUSION

Our study in collaboration with the University Hospital of Lille allowed us to explore different algorithmic and statistical methods to analyze pharmaceutical prescriptions, with the aim of limiting drug iatrogeny. For this work, we used the MIMIC-III database ('Medical Information Mart for Intensive Care'). Three approaches have been proposed.

The first one consists in clustering methods, allowing the analysis of the prescriptions to be tested. This method has proven to be moderately effective. We have several avenues of improvement to enhance our clustering methods, including calculating distances between prescriptions differently, and testing the algorithms on other negative data.

The second approach consists in supervised learning. Machine learning methods have given excellent results and have allowed us to discern the most important features in the prescription classification process.

The third approach consists in identifying potential drug interactions between two drugs prescribed to the same patient during the same hospital stay. To do this, we check whether or not these two drugs have been prescribed together before in our database.

Concerning the future perspectives, it would be interesting to explore the tracks explained before to improve the clustering methods. It would also be interesting to test our methods of analysis of potential drug interactions.

Acknowledgment

The authors thank XXX for their help in XXX.

References

[1] G. Nissim et T. Simon, The future of labor unions in the age of automation and at the dawn of AI , Technol. Soc., vol. 67, p. 101732, nov. 2021, doi: 10.1016/j.techsoc.2021.101732.

[2] F. Flamme-Obry et al., Impact du pharmacien clinicien sur la iatrogénie médicamenteuse chez le patient greffé rénal , Néphrologie Thérapeutique, vol. 14, no 2, p. 91-98, avr. 2018, doi: 10.1016/j.nephro.2017.04.004.

[3] W. Galanter, S. Falck, M. Burns, M. Laragh, et B. L. Lambert, Indication-based prescribing prevents wrong-patient medication errors in computerized provider order entry (CPOE) , J. Am. Med. Inform. Assoc., vol. 20, no 3, p. 477-481, mai 2013, doi: 10.1136/amiajnl-2012-001555.

[4] G. Reiner, S. L. Pierce, et J. Flynn, Wrong drug and wrong dose dispensing errors identified in pharmacist professional liability claims , J. Am. Pharm. Assoc., vol. 60, no 5, p. e50-e56, sept. 2020, doi: 10.1016/j.japh.2020.02.027.

[5] B. Aldughayfiq et S. Sampalli, A framework to lower the risk of medication prescribing and dispensing errors: A usability study of an NFC-based mobile application , Int. J. Med. Inf., vol. 153, p. 104509, sept. 2021, doi: 10.1016/j.ijmedinf.2021.104509.

[6] R. N. Keers, S. D. Williams, J. Cooke, et D. M. Ashcroft, Causes of Medication Administration Errors in Hospitals: a Systematic Review of Quantitative and Qualitative Evidence , Drug Saf., vol. 36, no 11, p. 1045-1067, nov. 2013, doi: 10.1007/s40264-013-0090-2.

[7] J. P. White, Health care, hospitals, and reengineering: The nightingales sing the blues , Good Jobs Bad Jobs No Jobs Transform. Work 21st Century Tor. Ont. Can. Harcourt Brace, 1997.

[8] H. Bouami et P. Millot, Healthcare Delivery System Security: the orchestration of automated and organizational solutions , IFAC-Pap., vol. 52, no 19, p. 127-132, janv. 2019, doi: 10.1016/j.ifacol.2019.12.123.

[9] E. Cousein et al., Effect of automated drug distribution systems on medication error rates in a short-stay geriatric unit , J. Eval. Clin. Pract., vol. 20, no 5, p. 678-684, 2014, doi: 10.1111/jep.12202.

[10] A. Yang et M. Grissinger, Wrong-patient medication errors: an analysis of event reports in Pennsylvania and strategies for prevention , Pa Patient Saf Advis, vol. 10, no 2, p. 41–9, 2013.

[11] Enquêtes Nationales sur les Événements Indésirables graves associés aux Soins - Comparaison des deux études ENEIS 2004 et 2009 — Direction de la recherche, des études, de l'évaluation et des statistiques . https://drees.solidarites-sante.gouv.fr/publications/documents-de-travail-1998-2019/enquetes-nationales-sur-les-evenements-indesirables (consulté le 11 février 2022).

[12] M. Veyrier, N. Bachalat, A.-L. Guenegou, B. Anne, P. Leglise, et D. Huchon Becel, Sécurisation de la prise en charge médicamenteuse des patients âgés pendant les permissions de sortie , Ann. Pharm. Fr., vol. 74, no 3, p. 212-221, mai 2016, doi: 10.1016/j.pharma.2015.10.001.

[13] C. des Comptes, Rapport Public 2014 . Dakar: Republic of Senegal, 2014.

[14] B. Allenet et al., De la dispensation au plan pharmaceutique personnalisé : vers un modèle intégratif de pharmacie clinique , Pharm. Hosp. Clin., vol. 54, no 1, p. 56-63, mars 2019, doi: 10.1016/j.phclin.2018.12.003.

[15] P. Bedouch et al., Trends in pharmacists' medication order review in French hospitals from 2006 to 2009: analysis of pharmacists' interventions from the Act-IP© website observatory , J. Clin. Pharm. Ther., vol. 40, no 1, p. 32-40, 2015, doi: 10.1111/jcpt.12214.

[16] E. Brudieu et al., Analyse pharmaceutique dans le cadre de la prescription informatisée , J. Pharm. Clin., vol. 18, no 3, p. 227-32, oct. 1999.

[17] A. H. Morice, S. Faruqi, C. E. Wright, R. Thompson, et J. M. Bland, Cough Hypersensitivity Syndrome: A Distinct Clinical Entity , Lung, vol. 189, no 1, p. 73-79, févr. 2011, doi: 10.1007/s00408-010-9272-1.

[18] J. Corny et al., A machine learning–based clinical decision support system to identify prescriptions with a high risk of medication error , J. Am. Med. Inform. Assoc., vol. 27, no 11, p. 1688-1694, nov. 2020, doi: 10.1093/jamia/ocaa154.

[19] Ruth Backman, Susan Bayliss, David Moore, Ian Litchfield, Clinical reminder alert fatigue in healthcare: a systematic literature review protocol using qualitative evidence , 2017 Dec 13;6(1):255. doi: 10.1186/s13643-017-0627-z

[20] C. Jarre et al., Analyse pharmaceutique ciblée des prescriptions à l'hôpital : évaluation d'une méthode de sélection des ordonnances , Ann. Pharm. Fr., vol. 75, no 6, p. 463-472, nov. 2017, doi: 10.1016/j.pharma.2017.07.001.

[21] F. Lagrange, J. Lagrange, C. Bénnaga, F. Taloub, M. Keddi, et B. Dumoulin, Surveillance médicamenteuse du sujet âgé : un outil de décision automatisé en pharmacie clinique , Pharm. Hosp. Clin., vol. 52, no 1, p. 10-20, mars 2017, doi: 10.1016/j.phclin.2016.11.003.