Predicting Medication Errors at Apollo Hospitals

Business Analytics & Intelligence (Batch 11)
Project Final Report

Submitted By:

Amrata Agrawal (BAI11005)

Harsh Upadhyay (BAI11015)

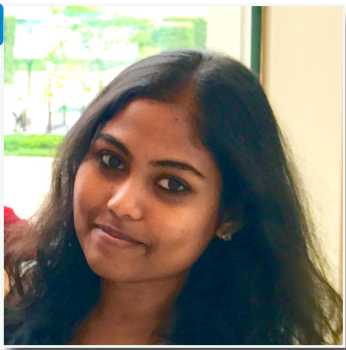
Jeya Sri R (BAI11017)

Sharan Sivakumar (BAI11047)

The Team

Amrata has experience in Data Warehousing and Business intelligence in healthcare domain. She has worked with various BI tools like Micro-strategy, Pentaho etc. to make real-time dashboards and reports for the clients. She is a Computer Science and Engineering graduate with PGDM in Finance. She has the knack of understanding things from business point of view and financial perspective. Her extensive research in food services in India was published and well acknowledged by the industry. She has an unexplained love for mathematics and would loves to solve mathematical puzzles like sudoku, Kakuro. She is the first Indian to register her name in "hall of fame" for solving SumSum puzzle.

Harsh enjoys solving intriguing problems and works with data and processes. He believes anything is complex till understood or explained and has a strong bias for action. An Instrumentation & Control Engineer with a PGDBA in finance, Harsh has an extensive experience of 17+ years in Business Optimisation, Data Science & Analytics and Digital Transformation. His expertise lies in reshaping the organisation level strategy by implementing disruptive innovation, steered cost saving initiatives through Digitalisation, Predictive & Prescriptive Analytics, Industry 4.0 and Informative Dashboards. As a deployment leader, he has created several mathematical models (time series, linear programming, logistic regression etc.) for solving the complex problem in Supply chain, Customer Life Cycle Management, Critical Business Decision on organisation burning issues. Harsh has spent 17+ years in various leadership roles true to his basic nature of learning and creating new things with companies such as LG Electronics, Motilal Oswal Financial Services, UnitedHealth Group & Bahrain Airport. Currently, he is leading the Business Transformation & Material Control Dept. in Lamprell Energy (Dubai, UAE). He is a certified Lean Six Sigma Master Black Belt, PMP certified & received a certification on Digital transformation course from MIT Sloan.



Inspired by programming and process improvement, **Jeya Sri** graduated with distinction in MBA and Computer Engineering. She has 4 years of diverse experience with work and people in Healthcare, Advertising, Media and Software industry. Her career path exposed her to consulting - product development, marketing and market research related projects. She also loves developing software products, participating in hackathons and equipping herself with open-source technologies. Being an avid book reader, she has an eye for details and reads between lines in professional workspace. Fascinated by data crunching, she is currently building a project portfolio that touch bases with Web scraping, Visualisation, Statistical Modelling and NLP.

An MBA gold medallist from Sri Sri University, Odisha, **Sharan** has more than 6 years of experience in varied fields such as fertiliser manufacturing, specialty chemical sales, dairy farming, and marketing of ITES for highly diverse industrial sectors. During these many years, he has undertaken various projects in process optimisation, digital marketing, new facility setup and management etc. The photographer and musician in him helps identify out of the box ideas to craft simple yet effective solutions to complex business problems. He has successfully completed a MOOC on Business Analytics from The Wharton School of Business, and also has obtained the Data Scientist badge from Tableau. Sharan is also a Kaggle enthusiast and loves coding Python.

Acknowledgements

We would first like to express our gratitude to Apollo Hospitals for sponsoring this project as a part of our Business Analytics & Intelligence program. We thank **Mr. Gaurav Loria,** Group Chief Quality Officer and VP Operations, Apollo Group for presenting us the vision and the bigger picture behind the problem statement and his guidance. **Dr. Mohammad Zia,** Clinical Pharmacologist, Apollo Hospitals, Jubilee Hills, Hyderabad and **Dr. Sanjeev Sharma,** Clinical Pharmacologist, Indraprastha Apollo Hospitals, Delhi, for all the long drawn discussions and their technical expertise that enabled us to dive deep into understanding the medication process at Apollo Hospitals. We thank **Ms. Yogamaya Nayak,** Senior Manager, Group Quality & Operations, Apollo Hospitals Group for her continued support, her relentless commitment to make this project a success as our coordinator from Apollo Hospitals.

We have been fortunate to receive the mentorship of **Dr. Haritha Saranga,** Professor - Production & Operations Management, IIM Bangalore, whose invaluable inputs have been the catalyst in surmounting the challenges posed by this project. We thank **Dr. U Dinesh Kumar,** Professor - Decision Sciences, IIM Bangalore & Programme Director, Business Analytics & Intelligence, for introducing us to this wonderful world of business analytics and his continued guidance. We also appreciate and thank **Ms. Pooja Balasubramanya,** Program Manager, IIM Bangalore for her support in helping us coordinate our project requirements with Apollo Hospitals.

Executive Summary

Premise

Apollo Hospitals

- Estd. in 1983, Apollo Hospitals is **India's leading healthcare provider.**
- 7000+ helping hands, 120 million lives touched.
- Encompasses over 10,000 beds across 70 hospitals, 4000+ pharmacies, 172 primary care & diagnostic clinics, 148 telemedicine units across 13 countries.
- **Pioneer in bringing state of the art, ground-breaking health care technologies to India**.

Medication Errors

- Ensuring that patients get their right medication is critical to quality.
- Medication errors usually occur during **prescription** of medicines; their **transcription** into databases; or **dispensing** at pharmacies.
- A wrongly **administered** medication *(drug/timing/dosage/mode)* to the patient can even prove to be fatal.

Project Aim

- **Employing data-driven decision making techniques to aid Apollo Hospitals to effectively minimise the number of medication errors and ultimately help save lives.**

Data

Collection & Evaluation

- Received medication error occurrences from 2017-2019 for 36 Indian hospitals across the Apollo group.
- Data columns included **error occurrence, type, sub-type & cause of error,** with the **staff & location** within the hospital.
- Homogeneity in the data capturing process was studied, and standardised data across the various hospitals was requested.

Quality Analysis & Cleaning

- Hospital data management not yet being automated & paperless can lead to a lot of challenges.
- **Improper date formatting, missing or junk data entries, duplicate records, mismatch of error codes with types, sub-types and causes** is very common and were cleaned.
- Collaborated with Apollo to identify the **top 6 hospitals** in terms of homogeneity in data.

Connecting Hospitals

- Since each hospital was unique in its operation, process improvement techniques such as **FMEA, FishBone, and RCA** were considered for implementation.

Exploration

Initial EDA

- The data had **only categorical variables**, and that meant that exploratory data analysis was carried out on **error counts, vis à vis error type, error subtype and error cause.**
- Monthly trends for error occurrences were also analysed, but since the date and time of errors were imprecise, and had formatting *(MM/DD/YY)* anomalies, these weren't used for analysis.
- Insights and patterns unique to certain specific hospitals, and those common among few/all hospitals were identified, categorised, and presented to Apollo.

Probing the Patterns

- A team of pharmacologists, and doctors were consulted to probe patterns in the insights.
- The **only process touchpoint involving electronic data capture** was determined, and this was the **indenting process.**
- As the indenting process was unique to a hospital, and the data also was voluminous, **one Group-A hospital with was zeroed-in.**
- Thus, the 36 hospitals were funnelled down to 1 prototype hospital with its own unique data capturing system.

Analysis

Solutioning

The Road Ahead

The Indenting Process

- The need for the indenting process arose, as all the data was only for errors, and not the day-day non-erroneous medication data.
- **Indenting is the process of entering into a database, the medications across all prescriptions to be dispensed on a periodic *(daily)* basis.**
- The indenting data received had **15 lakh records**, across 5 features *(DrugCode, Ward, GenericName, Dose, Mode)*
- **After cleaning, anomaly analysis & feedback from Apollo, and a final EDA, the errors in the error sheet were mapped against their respective indents.**
- A few errors were dropped, since they occurred much later, after indenting, during the discharge process.

Analytical Techniques

- None of the classification algorithms (*Naive Bayes, Logistic Regression, Decision Trees etc.)* that were tried gave comprehensible results, mainly since the **data simply lacked sufficient number of features to explain the variability in the errors.**
- Thus, **unsupervised learning algorithms** such as **Apriori & Clustering** were employed to group drugs, wards, and modes, based on error causing similarities.

Looking Beyond Indents

- The indent data had far too many records and far too less features, purely categorical.
- **Apriori algorithm helped determine the support and confidence between drugs while causing errors.** Support thresholds were used to filter out drug combinations causing errors.
- Since there is a possibility of **drugs prescribed in the future not being captured** in the 15 lakh indents, **drug related data from Apollo Pharmacy's website was scraped** to form a database. **Apollo Hospitals' official drug formulary** was also used to complement this database.
- The scraped data was further mapped with the results of EDA & Apriori to form an alert-based DSS.

Decision Support System

- Owing to limitations in the richness and quality of the indenting data, a DSS using **Natural Language Processing (NLP) & Text Mining / Analytics** for alerting possible future errors was designed, instead of a prediction model.
- **This decision support system would help raise flags to alert process owners (doctors, nurses, etc.), to probable future errors, based on any given combination of medications in an indent being entered in the future.**

The 5 Business Rules

- The alert system works on **5 hierarchical levels of business rules** that when triggered, will notify the user to potential risks.
- The **first rule** is generated with the **help of all patterns and insights from EDA**, contributing to higher error possibilities.
- The **second rule** is generated with insights from the **Apriori algorithm**, which further narrows down **drugs combinations** causing errors.
- The **third rule,** generated from the scraped database, accounts for **therapeutic drug substitutions.**
- The **fourth rule**, generated from the scraped database, accounts for **interactions between drugs**.
- The **fifth rule**, generated from the scraped database, accounts for **drug related contraindications.** (allergies, preconditions, etc)

Expanding the Horizons

- This alert system would help Apollo Hospitals to better identify potential medication errors, and act before the error becomes too costly.
- **This alert system would also serve the foundation stone in the digital transformation of the entire medication process, which, with sufficient rich data can enable far advanced AI to provide better healthcare, and save lives.**

Table of Contents

(click on page numbers to reach section)

Business Process Overview 7
- Hospitals in Today's Society 7
- The Medication Process at Apollo Hospitals 7

Problem Statement 8
- Medication Errors 8
- Apollo Hospitals' Key Challenges 8
- The Necessity for A Data-Driven Solution 8

Data Collection & Quality Analysis 9
- Error Occurrence Data 9
- Data Quality Analysis 9

Initial Exploratory Data Analysis 13
- Objectives of Exploratory Data Analysis 13
- Key insights Common Across All Hospitals 13
- Sample Graphs from the Initial EDA for the 6 Marquee Hospitals 14

Probing the Patterns & Process Door Analysis 16
- Mapping the Medication Process 16
- Factors Behind Prescription Errors - Fish Bone Analysis 19
- Risk Assessment Exercise & Priority Matrix 20

Working With Indenting Data 21
- Indenting Data Collection & Quality Analysis 21
- Mapping the Errors with the Indents 22
- Indenting Anomalies 23

Feature Engineering on Indenting Data 25
- The Need for Feature Engineering 25
- Information Retrieval (IR) Process for Text Mining & Text Analytics 27

Final Exploratory Data Analysis 30
- Key EDA insights from Indents 30
- Key EDA insights from Errors 32

Dimensionality Reduction & Classification - Supervised Learning Models 35

Building A Data-Driven Decision Support System 36
- Structure of the DSS 36
- DSS Functions & The Business Rules 37
- Monitoring & Reporting Alerts with the DSS Dashboards 37
- Creating a Knowledge Base with Natural Language Processing for the Business Rules 42

Benefits & Caveats of the Decision Support System 49
- Benefits of the DSS 49
- Caveats of the DSS 50

Conclusion & Recommendations 51
- Proposed Solutions 51
- Recommendations 53

Bibliography 54

Business Process Overview

Hospitals in Today's Society

Hospitals mark important central points in people's lives, from the birth of an offspring to those on the verge of death, and having quality hospitals around matters a lot to us. Hospitals reflect the needs and values of the communities in and around them, while also being resilient and able to maintain and scale up services in emergency situations. Effective hospitals such as Apollo Hospitals are designed for their users, with attention to the needs of special populations, such as children and the elderly. A well-designed hospital environment maximises the effectiveness of clinical care delivery and enhances the well-being of patients and hospital staff.

Hospital functions and organisation vary according to health-care delivery organisations and each hospital's unique position in the system. Good management structures ensure coordination among staff, services, infrastructure and supply chains to deliver high-quality care. Clinical registries and structured data audits facilitate rapid identification of high-yield areas for improvement. Regular monitoring of service quality with targeted intervention to address gaps drives ongoing improvement.

The Medication Process at Apollo Hospitals

For the last four decades, Apollo Hospitals has redefined healthcare standards in India, with groundbreaking strides in improving patient safety & experience. A network of healthcare establishments with 10000+ beds, 7000+ doctors, 4000+ pharmacies is both vast, and complex. Every Apollo hospital unit is a sophisticated system where time is of utmost essence, with a multitude of simultaneous processes that imbricate, are interconnected and are interdependent to achieve the highest standards of quality.

One such process critical to ensure the safety of inpatients is the medication process. The process begins at the doctor's consultation room, or at an emergency ward, where doctors diagnose the disorder and prescribe an appropriate set of medicines. These medicines vary in strength of drug, dosage, mode of administration, duration of the course, times to be administered in a day, and so on.

The prescriptions, after a prescription level audit, enter the transcription process, where the drugs needed by every patient is entered into an indenting database, on a periodic basis. This could be every 12, 24, or 48 hours, and varies from one hospital unit to another.

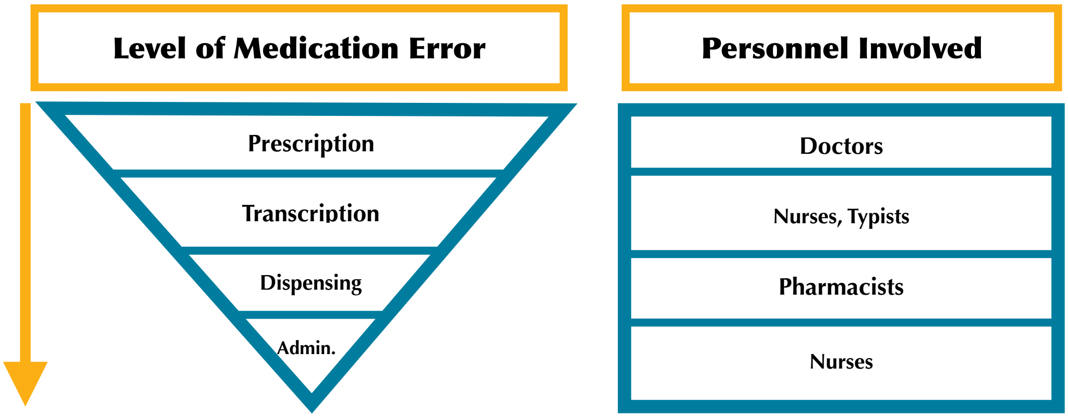
The indented drugs are accessed by the inpatient pharmacies, which dispense the drugs for the respective patient, and the dispensed medication are ultimately administered to the patient by the nurse in charge.

Unlike outpatients, inpatients also have routine visits from duty doctors, wherein prescriptions could be altered based on the patient's prognosis, until the time of discharge, where a new set of medication are freshly prescribed.

And a medication error could happen anywhere in this process.

Problem Statement

Medication Errors



A medication error is any preventable event that may cause or lead to inappropriate medication use or patient harm while the medication is in the control of the health care professional. Such events may be related to professional practice, health care products, procedures, and systems, including prescribing, product labelling, packaging and nomenclature, dispensing, administration, education, monitoring, and use.

Apollo Hospitals' Key Challenges

Monitoring the multifold reasons behind medication errors can be very cumbersome. The eight most common causes behind medication errors are communication problems, inadequate information flow, patient related issues, human error, organisational transfer of knowledge, staffing patterns and workflow, inadequate policies, and technical failures.

Apollo Hospitals desired to minimise the total number of medication errors in their hospitals, but most importantly, ensure that no error reached the administration level, i.e. reach the patient to cause harm. The harm a medication error could cause to a patient ranges from minor symptoms to potentially even death.

The Necessity for a Data-Driven Solution

The way forward to achieving this goal of preventing errors from reaching patients was to effectively understand the sources of such errors, their direct causes, and the indirect factors influencing them. A cursory exploration of Apollo's medication error data across 36 Indian hospitals indicates that more than 80% of errors happened at the prescription stage, and a majority of the rest trickling down to the subsequent stages. Thus it became pertinent to focus on reducing prescription errors with data-driven-models, and to tighten the subsequent processes, reducing the snowballing of medication errors over time.

Apollo Hospitals understood that there was no scientific system currently in place to achieve this, and that a predictive analytics model would go a long way in preempting future medication error occurrences. This would not only result in the streamlining of processes to enhance the overall productivity of the system, but also directly impact Apollo Hospitals' mission of achievement and maintenance of excellence in healthcare for the benefit of humanity.

Data Collection & Quality Analysis

Error Occurrence Data

The initial datasets received were a collection of 36 files, comprising of error occurrence information from 36 hospitals across India. Apollo Hospitals units have been classified under three groups:

Group A: Hospitals catering to major metropolitan cities *(usually Tier-I cities/state capitals)* with the largest bed counts and the widest range of specialities.

Group B: Hospitals catering to smaller metropolitan cities (usually Tier-II cities) having slightly smaller bed counts than those in Group A.

Group C: Hospitals in large cities providing a niche speciality (*like cancer research centres or for children's / women's health)* or in much smaller cities with fewer beds than Group A & B.

The metadata for the 8 Group A, 10 Group B and 18 Group C hospitals are as follows:

Variable	DataType	Description
Date	Ordinal	Date when error is captured in the system
Hospital Name	Categorical	Name of the Hospital
City	Categorical	Hospital's location
Error Type	Categorical	Prescription, Transcription, Dispensing, Administration
Error SubType	Categorical	26 predefined subtypes for the 4 error types
Error Cause	Categorical	20 predefined causes for the errors
Drug Category	Categorical	Therapeutic / Pharmacological category of the drugs
Specialty	Categorical	Medical specialty under which patient is initially admitted
Ward	Categorical	Area inside the hospital where error occurs
Staff	Categorical	Staff involved in the error
Staff Role	Categorical	Consultant / Resident

Data Quality Analysis

The data provided in the raw dataset has been manually compiled from the various departments and wards in every Apollo Hospital unit, which in turn has been manually captured by the personnel to whom an error is reported.

Owing to the manual nature of data capture, there have been a vast number of data quality issues, and discrepancies in the dataset received. This also meant there was very little similarity between the procedures and methodologies to data collection in each hospital unit, resulting in unstandardised information being collected across the 36 units. A few of the data issues could be rectified immediately, but there were still many that needed the assistance of Apollo Hospitals for rectification.

A few of the many data anomalies have been listed below with examples:

Erroneous Date Formatting:

Interchange of MM and DD in the date format was present in many hospitals' data. For example, in the image below, the first two rows would be generally interpreted as 7th and 9th of January, but in reality they were found to be 1st of June and September, respectively

Time/Date/Month /Year	Error Type (prescription/transcription/ adminstration/dispensing)	Error Sub-Type (pick from the error sub-type map or mention sub-type code)	Error Cause (pick from the error cause map or mention cause code)	Mention Drug Category	Mention specialty- under which patient admitted
2019-07-01	ADMINISTRATION	E26	A5	ANTIEPILEPTIC	GASTROENTEROLOGY
2019-09-01	ADMINISTRATION	E14	A2	ANALGESICS	CARDIOLOGY
15/1/2019	ADMINISTRATION	E6	A2	ANTIEPILEPTIC	NEUROLOGY
18/1/2019	ADMINISTRATION	E26	A1	CEPHALOSPORINS	ORTHOPEDIC
20/1/2019	ADMINISTRATION	E26	A1	QUINOLONES	NEUROLOGY

Missing Data:

There were a lot of records missing across multiple sheets, like shown below:

Time/Date/ Month/Year	Error Type (prescription/transcription/ adminstration/dispensing)	Error Sub-Type (pick from the error sub-type map or mention sub-type code)	Error Cause (pick from the error cause map or mention cause code)	Mention Drug Category
05-01-2019	ADMINISTRATION	E11	A4	ANTITUBERCULAR
03-01-2019	PRESCRIPTION	E21	P10 (FREQUENCY	ANTIBIOTIC
05-01-2019	PRESCRIPTION	E5	P2	ANTIBIOTIC
07-01-2019	PRESCRIPTION	E6	P9	XANTHINE OXIDASE INHIBITOR
08-01-2019	PRESCRIPTION	E2	P9	LMWH
11-01-2019	PRESCRIPTION		P4	ANTIHYPERTENSIVE STATINS

Junk Data:

Fields supposed to capture a specific variable had unintelligible data, or data belonging to another field. As shown below, instead of specialty of admission, random numbers were found:

Error Sub-Type (pick from the error sub-type map or mention sub-type code)	Error Cause (pick from the error cause map or mention cause code)	Mention Drug Category	Mention specialty-under which patient admitted
E21	T8	ANTIBIOTICS	159084
E21	T8	ANTIBIOTICS	158540
E17	T2	ANTINEOPLASTIC	158556
E21	T8	ANTIPLATELET	159640
E2	T9(NURSE INDENT WITHOUT PROPER REFERRING TO DRUG CHART)	ANTINEOPLASTIC	158457
E19	T8	ANTIGOUT	158131
E1	T8	HORMONE	158708
E20	T9 (NURSE NOT AWARE ABOUT THE DRUG GENERIC)	ANTIPLATELET	159519
E21	T8	ANTIBIOTICS	159316
E21	T8	ANTINEOPLASTIC	160547

Data Duplication:

Error subtypes and causes assigned specifically for one error type (say Prescription) overlapped with another, leading to ambiguity about the errors being mutually exclusive, i.e. whether the same error was being doubly or triply captured at multiple stages. As shown in the image below, Errors 594 & 595 seem like they are both the same error, but it is impossible to determine whether it is a prescription error, or an administration error, and if it occurred once or twice.

S.no.	Time/Date/Month/Year	Error Type (prescription/transcription/ adminstration/dispensing)	Error Sub-Type (pick from the error sub-type map or mention sub-type code)	Error Cause (pick from the error cause map or mention cause code)	Mention Drug Category	Mention specialty-under which patient admitted	Name of the unit/area where error occured	Mention staff-type involved in the error (Dr/Nurse/Pharmacist)	For staff-type Dr-mention the role(Resident/ Registrar/DNB/ Consultant)
Error 59	29.03.20,10:45AM	ADMINISTRATION	E-26 (DOSE, NOT ADMINISTERD TO THE PATIENT)	A-1	ANTIPSYCHOTIC	NEPHROLOGY	6TH FLOOR	NURSE	
Error 159	8.1.20,5.15PM	ADMINISTRATION	E-21	A-5(WRONG TIMINGS)	ANTIBIOTICS	Internal Medicine	5TH FLOOR	NURSE	
Error 259	6.12.19/3:00PM	ADMINISTRATION	E26(DOSE NOT ON TIME)	A-5(STAFF WAS BUSY)	ANTI-DIABETIC	Internal Medicine	6TH FLOOR	NURSE	
Error 359	14.11.19,8:30AM	ADMINISTRATION	E-21	A-1	ANTIPYRATIC	Internal Medicine	ICU-7	NURSE	
Error 459	30-10-19,12:15PM	PRESCRIPTION	E-26 (SOS PLACE)	P-1	BENZODIAZEPINES	Nephrology	ICU-1	DOCTOR	RESIDENT
Error 559	8-8-2019 8.20AM	PRESCRIPTION	E-26(SOS)	P-1	ANTIEMETIC	Medical Oncology	MICU	DOCTOR	CONSULTANT
Error 590	27-08-19 5.35PM	PRESCRIPTION	E-26(HIGH DOSE)	P-1	ANTIHYPERTENSIVE	Neurosurgery	ICU-6	DOCTOR	CONSULTANT
Error 591	27-08-2019 5.15PM	PRESCRIPTION	E-21	P-2	ANTIFUNGLE	Paediatrics	ICU-7	DOCTOR	CONSULTANT
Error 592	28-08-19 4.55pm	TRANSCRIPTION	E-21	A-1	THYROID DEFICIENCY	Surgical Gastroenterology	ER	NURSE	
Error 593	28-08-19 5.05PM	TRANSCRIPTION	E-21	P-2	ALPHA ADRENERGIC ANTAGONIST	Urology	ER	DOCTOR	CONSULTANT
Error 594	29-08-19 4.45pm	PRESCRIPTION	E-26	P-1	NSAIDS	Neurosurgery	ER	DOCTOR	RESIDENT
Error 595	29-08-19 3.45PM	ADMINISTRATION	E-26	A-4	NSAIDS	Neurosurgery	ER	Nurse	
Error 596	29-08-2019 5.15PM	PRESCRIPTION	E-10	A-1	DUST	Internal Medicine	ICU-6	DOCTOR	CONSULTANT
Error 597	29-08-2019 5.35PPM	PRESCRIPTION	E-19	P-1	ANTIHYPERTENSIVE	Internal Medicine	ICU-6	DOCTOR	CONSULTANT
Error 598	29-08-19 5.05PM	TRANSCRIPTION	E-21	P-2	ANTIHYPERTENSIVE	Neurology	ICU-2	DOCTOR	CONSULTANT
Error 599	30-08-2019 09:45	ADMINSTRATION	E-21	A-4	ANTIBIOTICS	Pulmonology	5TH FLOOR	NURSE	
Error 659	7-6-2019 12.24PM	ADMINSTRATION	E-20	A-5(NOT AWARE ABOUT BRANDS)	ANTIBIOTICS	Pulmonology	ER-OBS	NURSE	
Error 759	03-5-19 10.02AM	PRESCRIPTION	E-10	P-1	ANTIBIOTICS	Surgical Oncology	6TH FLOOR	DOCTOR	RESIDENT

Data Mismatch - Excess Error SubType Codes:

26 error subtypes were defined across all 4 error types, but there were some hospitals which exceeded this, as shown below:

Time/Date/Month /Year	Error Type (prescription/transcription/ adminstration/dispensing)	Error Sub-Type (pick from the error sub-type map or mention sub-type code)	Error Cause (pick from the error cause map or mention cause code)	Mention Drug Category	Mention specialty-under which patient admitted
26-02-2018	prescription error	E26	P9	HYPOTHYROID	OBG
12-02-2018	prescription error	E26	P9	VITAMINS	OBG
18-02-2018	prescription error	E27	P9	PROTON PUMP INHIBITORS	ENT

Data Mismatch - Error Types vs Error SubTypes vs Error Cause:

In a few hospitals, descriptions of errors were given, instead of the predefined codes.

Time/Date/Month/ Year	Error Type (prescription/transcription/ adminstration/dispensing)	Error Sub-Type (pick from the error sub-type map or mention sub-type code)	Error Cause (pick from the error cause map or mention cause code)	Mention Drug Category
09.01.17	Transcription error	Wrong patient	Forgetfulness/hurry	Regular drugs
10.01.17	Transcription error	Excess quantity indented.	Forgetfulness/hurry	Antibiotic
15.01.17	Transcription error	Wrong frequency or wrong regimen	Lack of awareness of writing the complete drug order while	Anti anginal
17.01.17	Transcription error	Wrong drug	Lack of awareness of writing the complete drug order while	Corticosteriod

Since these issues as highlighted above were prevalent throughout most of the hospitals; there was a need for standardised data to compare errors across multiple hospitals, and it was cumbersome for clean data across all 35 hospital units, Apollo Hospitals was presented with two alternatives:

i. Identify a geographical region having a combination of Groups A,B, and C hospitals, which were handled by the same regional quality team and should hence have standardised data collection procedures, or

ii. Identify a few marquee hospitals across the country that had relatively better data collected, and work on fixing the error data collection in those hospitals.

Apollo Hospitals chose the second option, and revised data for error occurrences from 2017-2019 was collected for 6 marquee hospital units across India, which were a combination of Groups A, B, and C.

Initial Exploratory Data Analysis

Objectives of Exploratory Data Analysis

Six marquee units were identified by Apollo Hospitals and post the receipt of cleaned error data for them, EDA was performed with the objective of determining **"what"** *(types of error, kind of drugs)*, **"when"** *(time of error occurrence)*, **"where"** *(ward / specialty)* and **"whom & how"** *(staff)* factors contributed to these errors, in order to prevent prescription, transcription, and dispensing errors from reaching the patients (administration errors).

Fields used in the EDA:

i. Date, Error-Type, Error-SubType, Error Cause, Drug Category, Specialty, Area, Staff Role
ii. Drug name, Dosage, Mode of Administration, Time
iii. Two additional unit-level information (Bed Count & Monthly discharges) were used for computing derived variables for standardising error counts across multiple hospitals.

Time Period: 2017-2019

Target Hospitals: 6 units *(Combination of Group-A & Group-B)* hospitals.

Since all the data fields were categorical, analysis conducted were on the count of errors, vis à vis Error-Type, Error Cause, and so on..

Key Insights Common Across all Hospital Units

i. **Error Type:** Prescription errors accounted for a majority of all errors occurred (~80%), and also had the highest year-on-year increase (134%).

ii. **Error SubType:** "Wrong Dose" (E19) ,"Wrong Drugs" (E20), "Dose not mentioned" (E8) "Documentation Error" (E6) & "Others" (E26) are the highest medication error sub-types, and their total contribution reached to 90%

iii. **Error Cause:** 77% of all errors are because of Forgetfulness/Hurry (P9 - Doctor Hurry & A4 - Nurse Hurry) specially in the night shift (57%)

iv. **Specialty:** Four medical specialties contributed to ~52% of medication errors.

v. **Drug Category:** 33% of all medical errors stem from just three drug categories namely, Antibiotics, Cardio-Vascular Drugs & Antihypertensives.

vi. **Criticality:** 97% errors are not related with High Alert Medication

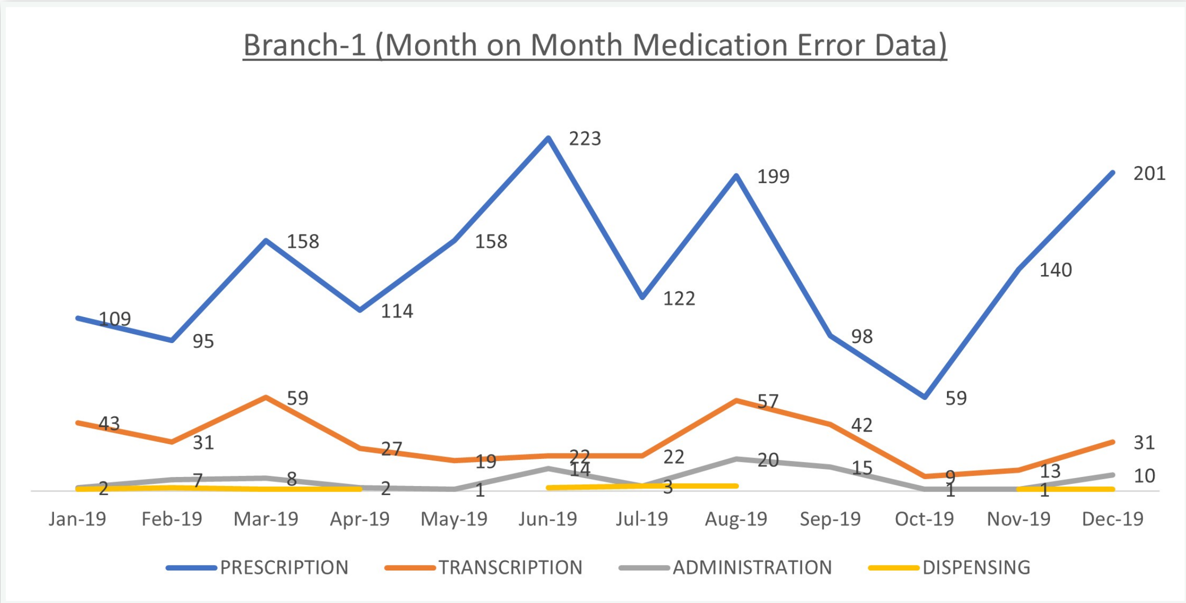
vii. **Ward:** 89% of errors are happened in Wards in comparison with ICUs.

viii. **Staff:** Doctors contributed to more errors (64%) in general, than nurses (27%), and the rest distributed among Pharmacists, Typists, Dieticians etc.

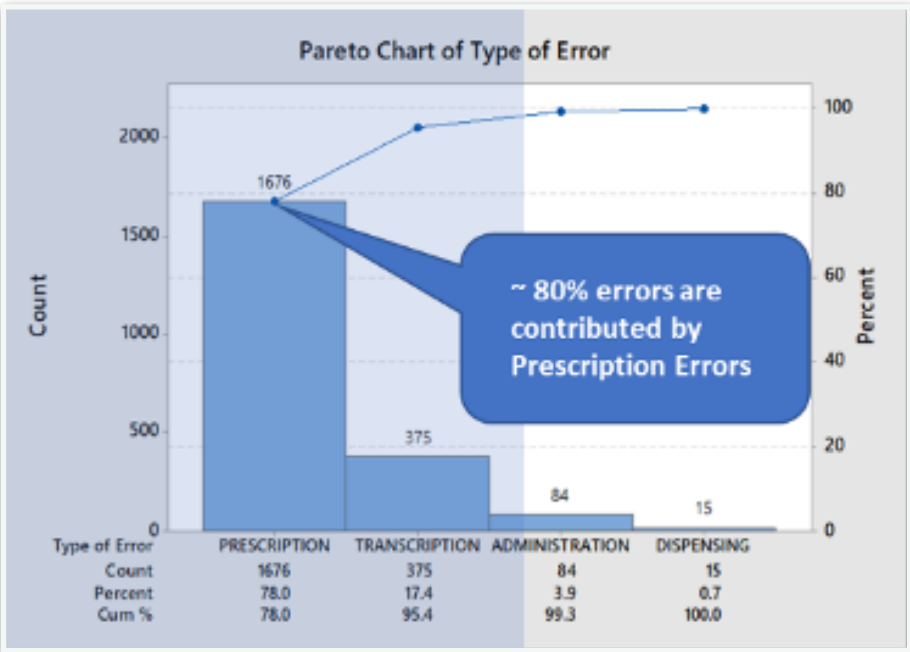
ix. **Timeline Trends:** Errors per 100 beds across hospitals showed different trends; while some hospitals had more errors in Q1 & Q2, others had more in Q3 & Q4.

Sample Graphs From the Initial EDA for the 6 Marquee Hospitals

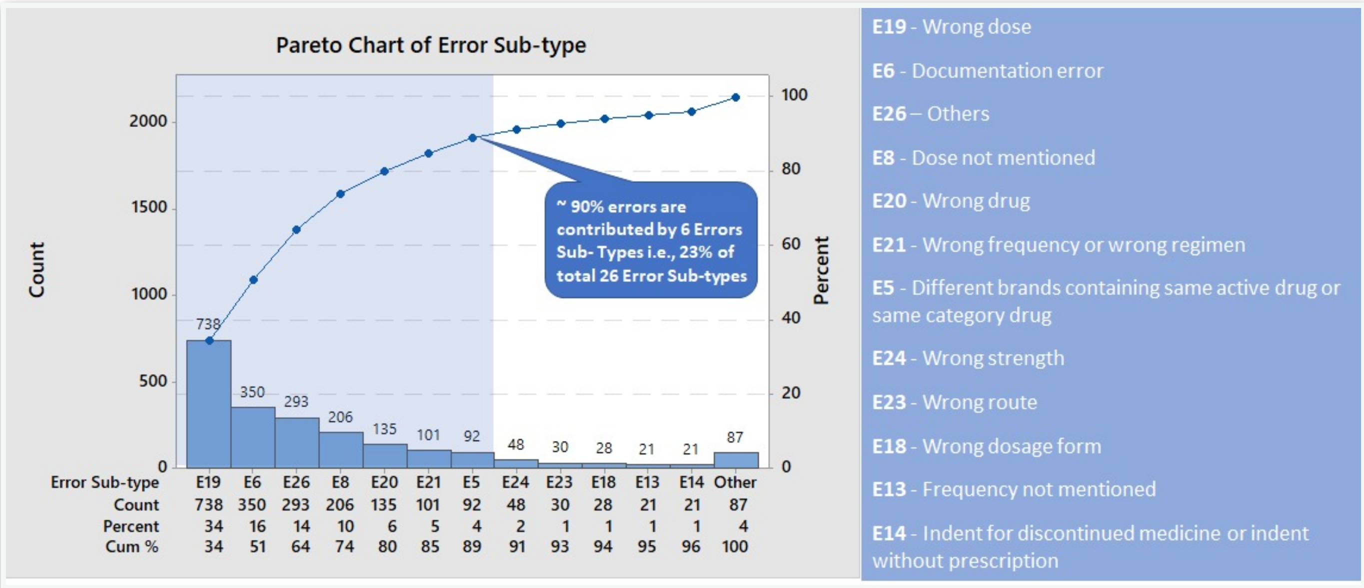
Month-Month Growth In Error Occurrences In A Group-B Hospital In 2019



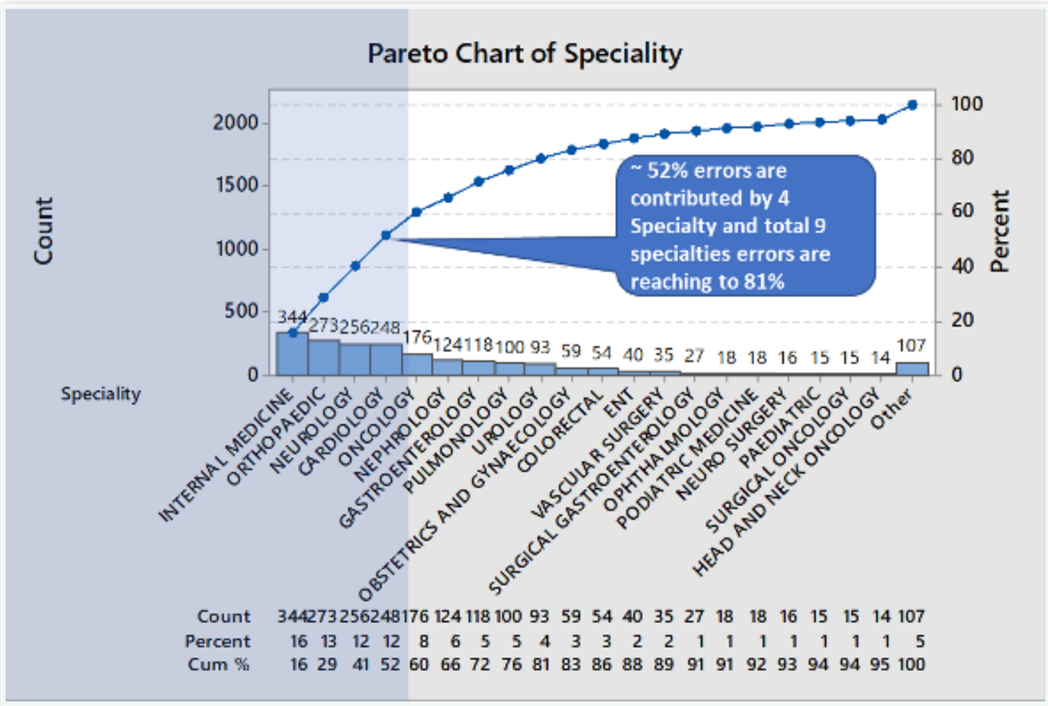
Pareto Chart Analysis for Error Types contributing to Errors Across Hospitals



Pareto Chart Analysis for Error SubTypes contributing to Errors Across Hospitals



Pareto Chart Analysis for Specialty contributing to Errors Across Hospitals



Probing the Patterns & Process Door Analysis

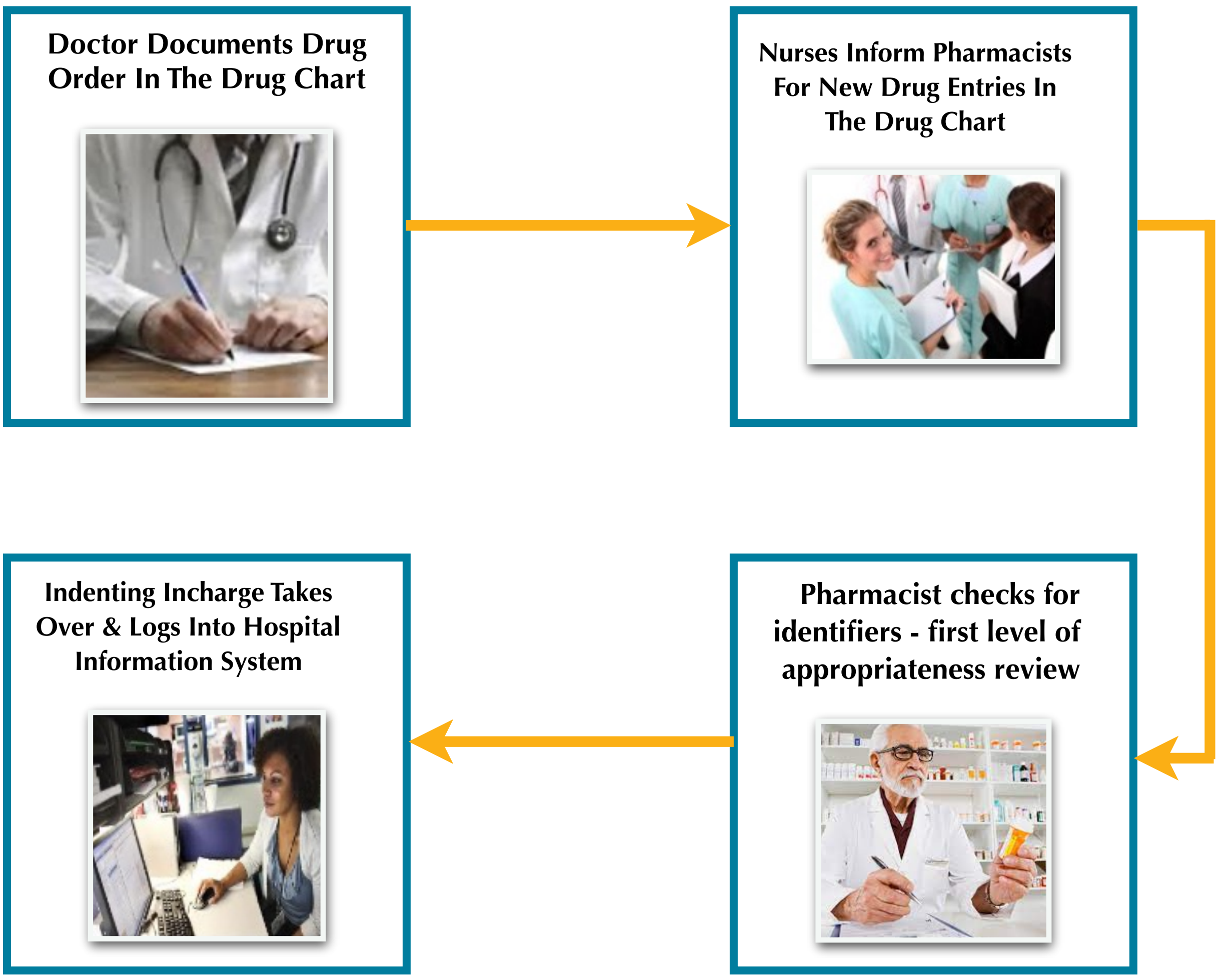
Since the insights from EDA posed more questions behind the "**why do medication errors occur in the first place"**, it was essential that these insights be shared with Apollo Hospitals, and then probed for a more detailed understanding of the process dynamics at each of these units and why these errors stemmed.

After several discussions with a team of experts comprising of clinical pharmacologists, emergency doctors, and members from Apollo's quality assurance team, a comprehensive understanding of the crux of the medication process at Apollo Hospitals was obtained.

This was followed by a few process improvement techniques, as listed below:

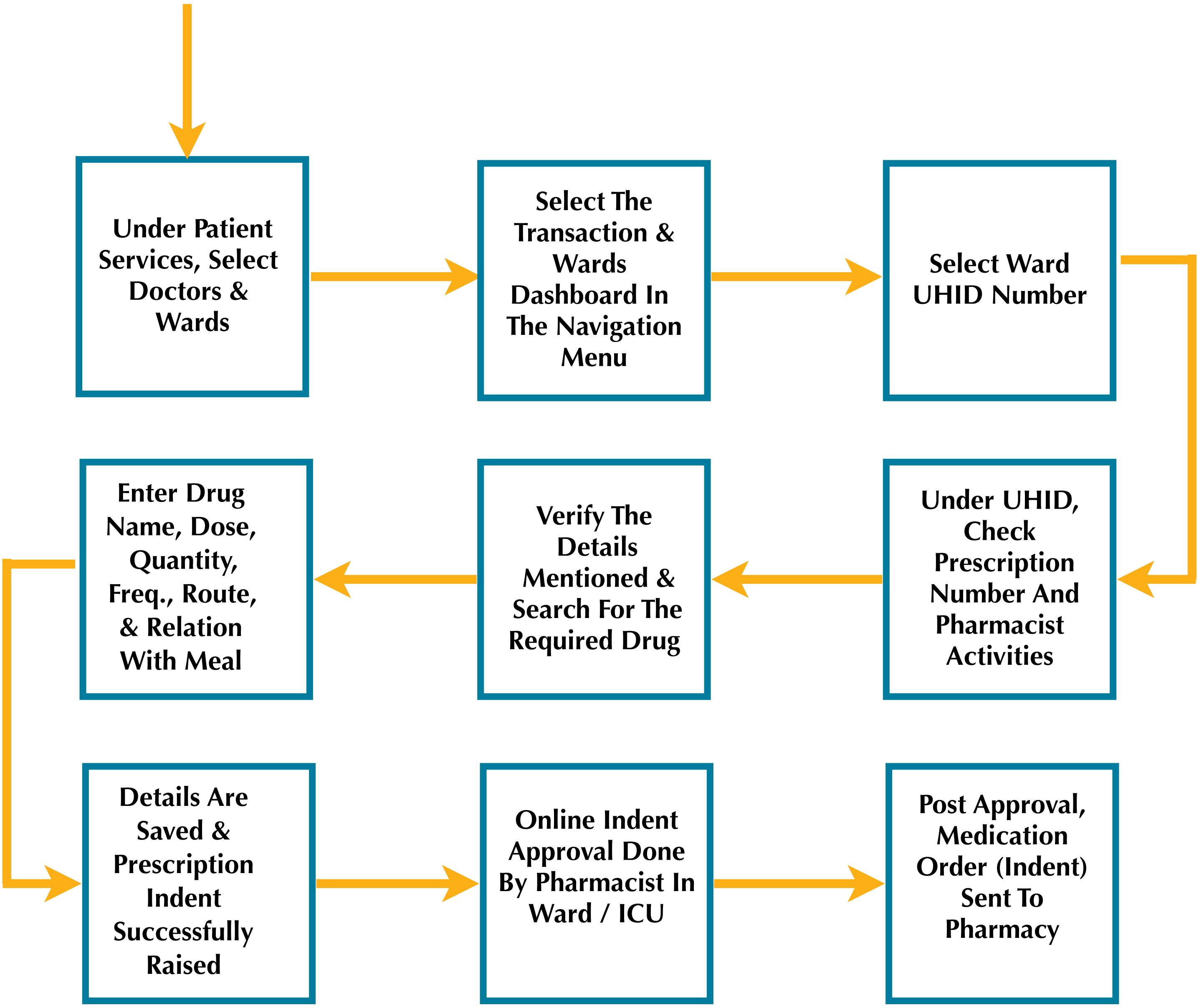
i. Detailed **process mapping** for the medication process,
ii. Categorisation of error-causing factors into buckets under **FishBone analysis**,
iii. Constructing FMEA to identify for the top 10 factors **risk assessment exercise**.
iv. Grouping the different critical factors to be addressed under **Process Door and Data Door Solutioning.**

Mapping the Prescription-Transcription-Dispensing Process for Medications

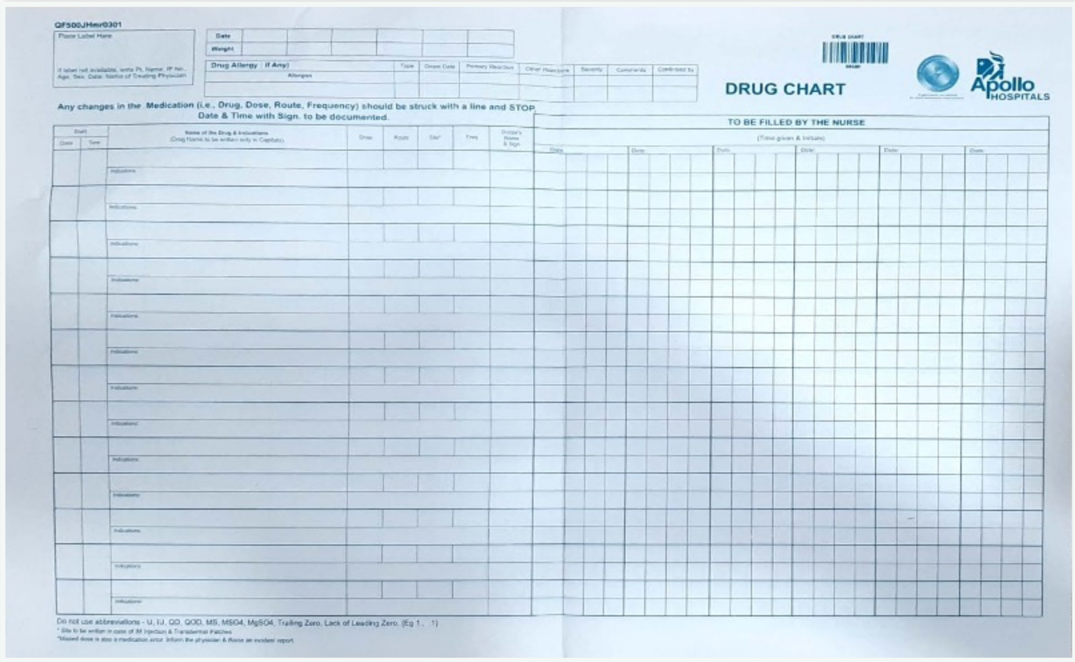


Mapping the Prescription-Transcription-Dispensing Process for Medications (contd..)

(The process continued below happens at the point of transcription *(indenting process in-charge)*)



The Apollo Hospitals Drug Chart



QF500JHmr0301

Place Label Here

If label not available, write Pt. Name, IP No., Age, Sex, Date, Name of Treating Physician

DRUG CHART

Date						
Weight						

Drug Allergy : If Any) Allergen	Type	Onset Date	Primary Reaction	Other Reactions	Severity	Comments	Confirmed by

Any changes in the Medication (i.e., Drug, Dose, Route, Frequency) should be struck with a line and STOP Date & Time with Sign. to be documented.

TO BE FILLED BY THE NURSE (Time given & Initials)

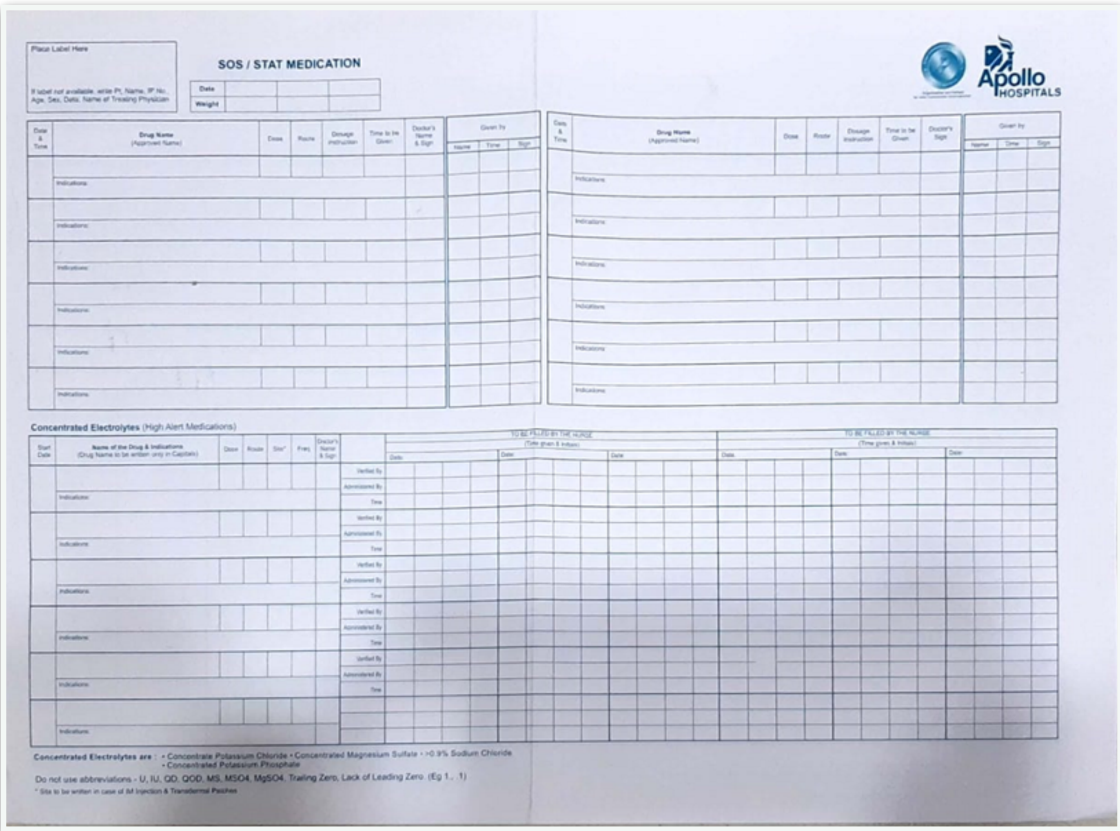
Start Date	Start Time	Name of the Drug & Indications (Drug Name to be written only in Capitals)	Dose	Route	Site*	Freq	Doctor's Name & Sign	Date	Date	Date	Date	Date	Date
		Indications:											

Do not use abbreviations - U, IU, QD, QOD, MS, MSO4, MgSO4, Trailing Zero, Lack of Leading Zero. (Eg 1., .1)

* Site to be written in case of IM Injection & Transdermal Patches

*Missed dose is also a medication error. Inform the physician & Raise an incident report.

The Apollo Hospitals SOS/STAT Medication Chart



Place Label Here

If label not available, write Pt. Name, IP No., Age, Sex, Date, Name of Treating Physician

SOS / STAT MEDICATION

Date			
Weight			

Date & Time	Drug Name (Approved Name)	Dose	Route	Dosage Instruction	Time to be Given	Doctor's Name & Sign	Given by Name	Given by Time	Given by Sign
	Indications:								

Concentrated Electrolytes (High Alert Medications)

TO BE FILLED BY THE NURSE (Time given & Initials)

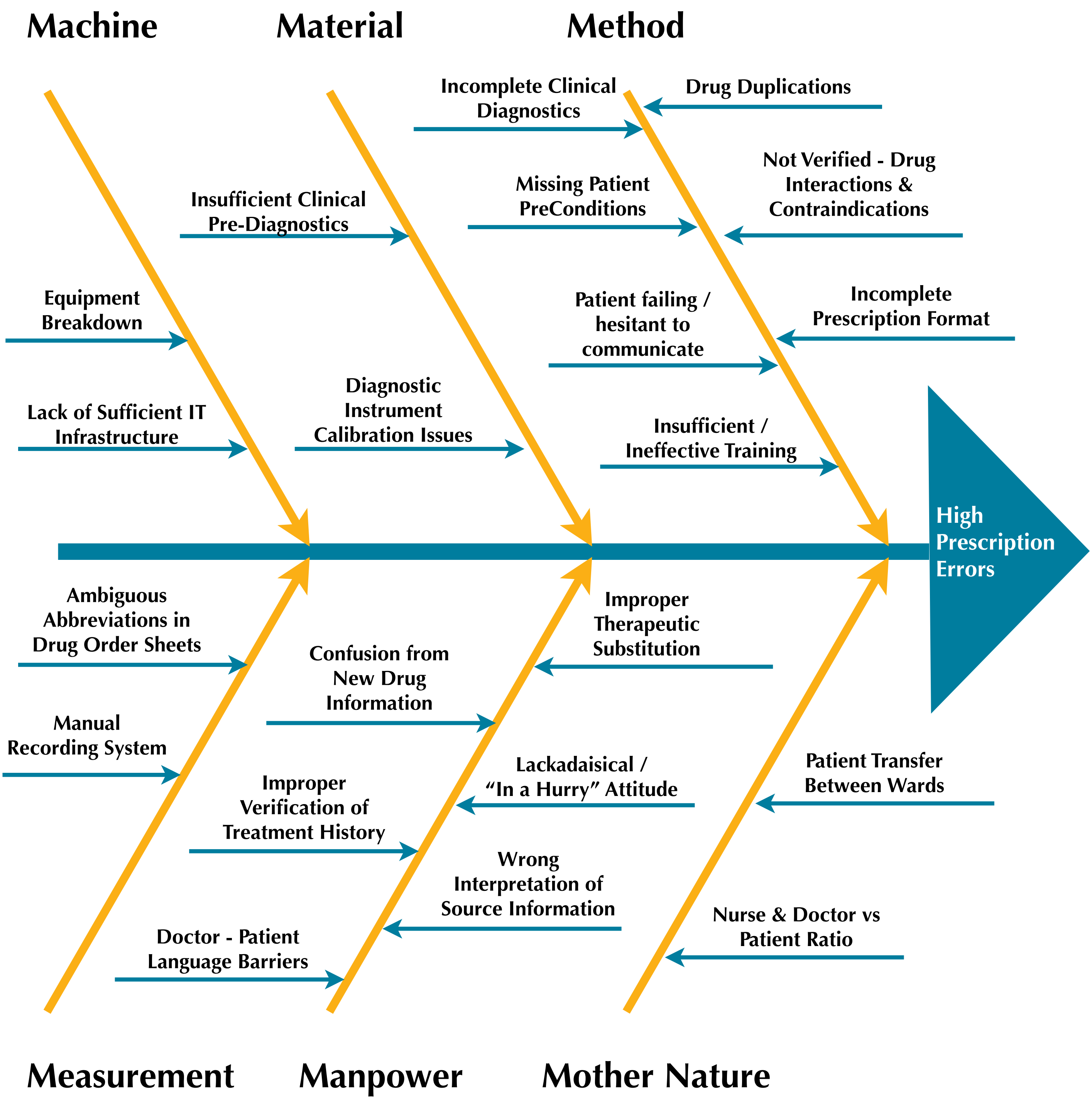
Start Date	Name of the Drug & Indications (Drug Name to be written only in Capitals)	Dose	Route	Site*	Freq	Doctor's Name & Sign		Date	Date	Date	Date	Date	Date
	Indications:						Verified By / Administered By / Time						

Concentrated Electrolytes are : • Concentrate Potassium Chloride • Concentrated Magnesium Sulfate • >0.9% Sodium Chloride • Concentrated Potassium Phosphate

Do not use abbreviations - U, IU, QD, QOD, MS, MSO4, MgSO4, Trailing Zero, Lack of Leading Zero. (Eg 1., .1)

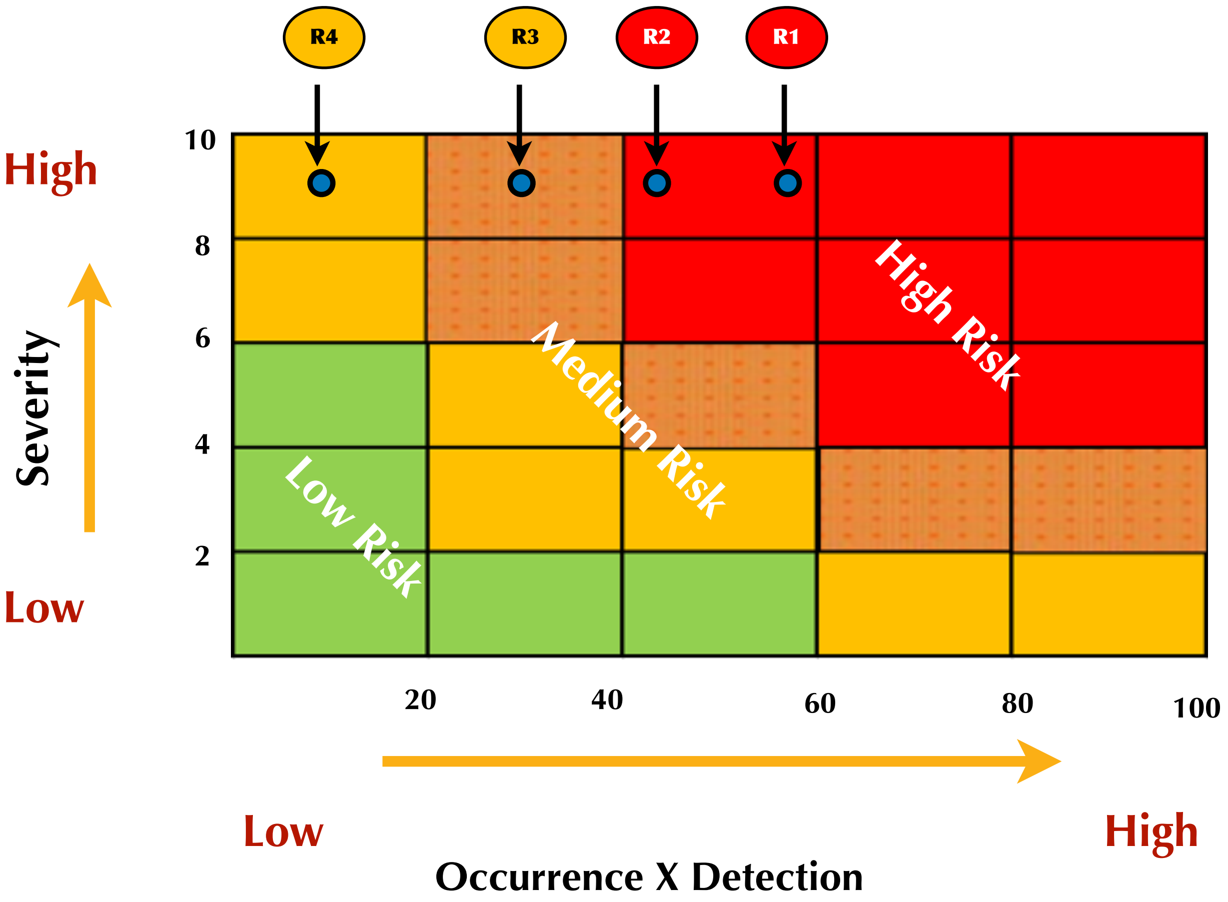
* Site to be written in case of IM Injection & Transdermal Patches

Categorisation of Prescription Error Factors Into Buckets Under FishBone Analysis



The various probable factors behind prescription errors were grouped under their respective buckets and the top few contributing factors selected for Risk Assessment Exercise, as shown below:

Risk Assessment Exercise & Priority Matrix



High Risk Priority Factors

R1: Drug Interactions & Contraindications Not Verified

R2: Drug Duplications

R3: Improper Therapeutic Substitution

R4: Confusion From New Drug Information

Of the top 10 high risk factors identified by FMEA, 6 were associated with process improvement, and were to be handled with the Process Door analysis techniques in place, by regional quality team of Apollo Hospitals. For the Data Door analysis, our focus was to develop a data-driven prediction-based model to incorporate the effects of these remaining top 4 factors.

Working with Indenting Data

The outcome of process door analysis directed us towards a few high critical factors that may affect the occurrence of medication errors. Till this instant, all the datasets received had only error occurrence information and did not have any data pertaining to day-to-day medication activities. That is, from a classification standpoint, only positive class data *(error occurrence)* had been provided.

Since most of the data being captured at Apollo hospitals' units were manual, the one process touchpoint were data was being captured digitally had to be identified. **That was the indenting process.**

Indenting Data Collection & Quality Analysis

The high volume of day-day medication issued at any moderately sized Apollo hospital unit ran into multiple lakhs per year. Despite having performed initial EDA for 6 marquee hospitals, it was prudent to perform the indenting data analysis for one hospital first, and then hopefully be able to replicate the process in other units as well.

The indenting dataset was received for 12 months, from one hospital, and it had more than 15 lakh records with the following variables:

Variable	DataType	Description
Created Date	Ordinal	Date when indent is raised
UHID	Categorical	Patient Unique ID
IP Number	Categorical	Inpatient Number (Unique to every admission)
Medicine Name	Categorical	BrandName of the Drug
Dosage	Categorical	Strength of the Drug
DrugCode	Categorical	Unique ID for each medication
Generic Name	Categorical	Generic Name of the Drug
Relationship with Meal	Categorical	Before / After Meals
Route of Administration	Categorical	Mode in which drug is given
WardName	Categorical	Location in the hospital

A few of the columns from this dataset represented similar information *(like DrugCode & Medicine Name)*, and a few fields did not have consistent or properly collected data like *(Dosage/Frequency/Relationship)* as they weren't applicable to most of the indented medication.

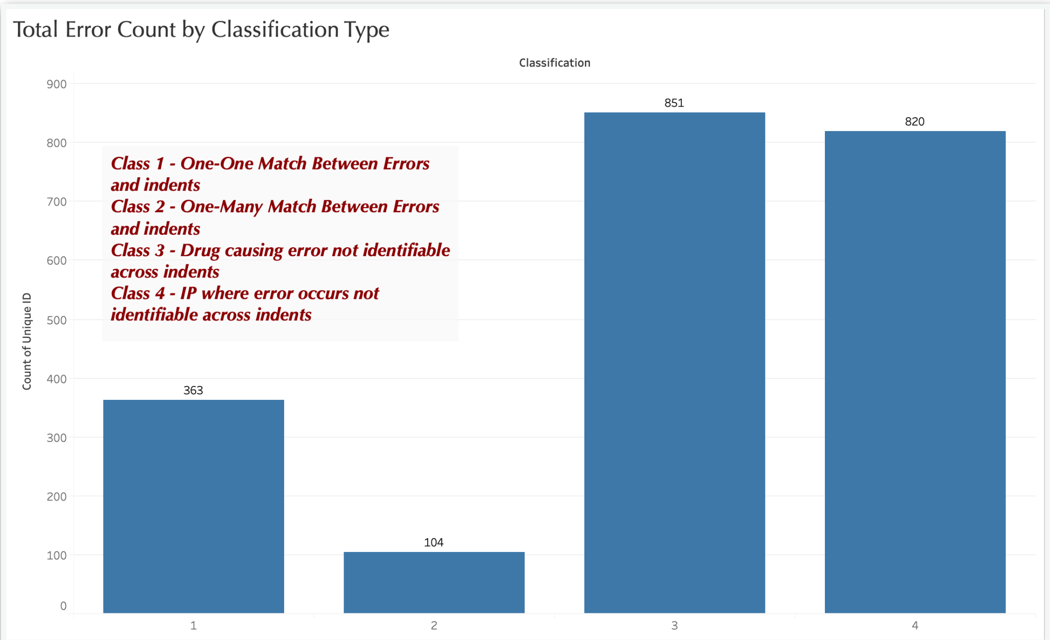
Mapping the Errors With the Indents

The next step in data pre-preprocessing was to the map all the 2150 errors that occurred from Jan-Dec 2019 in this hospital unit with their respective records present among the 15 lakh indents.

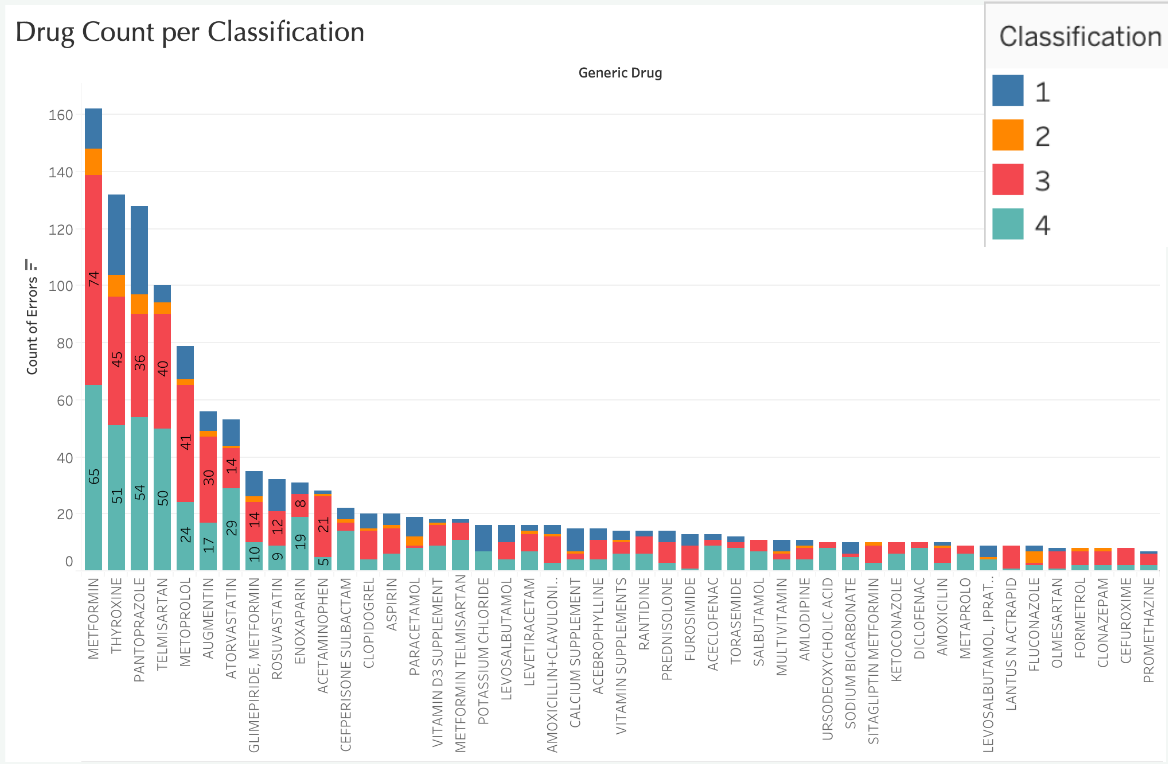
The following methodology was applied for classification of all the indents:

Category	Description
0	Indented medication did NOT cause any medication error
1	The error present in the error sheet has a 1-1 match in the indent sheet
2	The error present in the error sheet has a 1-many match in the indent sheet
3	Error sheet record has a IP match, but erroneous drug is missing in the indent sheet
4	Error sheet record does NOT have IP match in the error sheet

This mapping resulted in the error distributions amongst the categories 1- 4 as shown below:



Detailed EDA was performed on the 4 categories of errors to ascertain any relationships with the other explanatory features. A sample graph of distribution of categories 1- 4 among various Drugs (Generic Names) is shown below:



Indenting Anomalies

It can be seen that a majority of the errors belonged to either category 3 or category 4. Upon investigation, this was attributed to the various anomalies present in the indenting data, of which the most common and critical ones are highlighted below *(IP & UHID Numbers have been blocked to preserve patient confidentiality)* **:**

Error Date and Indent Date Mismatch

Error Sheet

	C	E	F	G	H	L	N	R
1	Column 1	Column 3	Column 4	Column 5	Column 6	Column 10	Column 12	ks (3/
1920	02-12-2019	E5	P9	ANTI ANGINAL	NEUROLOGY	ISOSORBIDE DINITRATE	[redacted]	4

Indent Sheet

F	H	I	S	T
CREATEDDAT	IPNUMBER	MEDICINENAME	DRUGCODE	GENERICNAME
25-12-2019	[redacted]	ATORVA 20MG TAB	ATO0042	ATORVASTATIN 20MG
25-12-2019	[redacted]	CILACAR 10MG TAB 10'S	CIL0012	CILNIDIPINE 10MG
25-12-2019	[redacted]	CILACAR 10MG TAB 10'S	CIL0012	CILNIDIPINE 10MG
25-12-2019	[redacted]	STROCIT 2 ML INJ	STR0003	CITICOLINE
25-12-2019	[redacted]	STROCIT 250MG INJ 4ML	STR0004	CITICOLINE
25-12-2019	[redacted]	ISMO - 10MG TAB 30'S	ISM0002	ISOSORBIDE MONONITRATE 10MG
25-12-2019	[redacted]	LOBET 20MG / 4ML INJ	LOB0015	LABETALOL
25-12-2019	[redacted]	LOBET 100MG TAB	LOB0014	LABETALOL 100MG

This is a typical scenario resulting in plenty of category 4 errors, where this particular patient has been admitted during the end of December 2019, with the indents confirming this, but an error has occurred on the 2nd December 2019. There were many such cases of perplexing date conflicts. Apollo Hospitals confirmed that this is a common practice since many patients choose to bypass the hospital's medication system for billing / other reasons.

Error Recorded for One Patient is Present in The Indent of Another Patient

Error Sheet

C	D	E	F	L	N	P	R
Column 1	Column 2	Column 3	Column 4	Column 10	Column 12	Rem[illegible]ks (3/5	
12-07-2019	PRESCRIPTION	E13	P9	TELMISARTAN	IP# X	Error 1187	3
12-07-2019	PRESCRIPTION	E8	P9	CEFIXME		Error 1188	3
20-07-2019	PRESCRIPTION	E26(DRUG NOT ADDED IN DISCHARGE SUMMARY)	P4	TORASEMIDE	IP# Y	Error 1216	3
23-07-2019	PRESCRIPTION	E6	P6	ACECLOFENAC	[redacted]	Error 1259	3
23-07-2019	PRESCRIPTION	E21	P9	VITAMIN D3 SUPPLEMENT		Error 1243	3

Another scenario commonly prevalent was that the error present for a specific patient in the error sheet, on a certain date *(**IP# Y**, on 20th July 2019)* was actually present under the indents of another patient, on another date *(**IP# X**, on 12th July 2019)*!

This seemed like a coincidence at first, but once there were many repeated cases of the same, it was evident this too was an anomaly that resulted in category 3 errors.

Far Too Few Indents for Patients

There were scenarios where a certain patient had only one or two drugs indented in a day.

Error 2090	43811IP325158	12	12/12/19	RESIDENT	SERTRALINE	NO	[redacted]
Error 2091	43811IP325309	12	12/12/19	RESIDENT	AMOXICILIN	NO	
Error 2092	43811IP322971	12	12/12/19	RESIDENT	PANTOPRAZOLE	NO	
Error 1954	43812IP322604	12	13/12/19	RESIDENT	METFORMIN	NO	

Therapeutic Substitution & Nomenclature Substitution

Many cases of category 3 errors were due to the fact that either:

i. Duty doctors had changed to a therapeutic substitute for a given drug, *(like Atorvastatin over Rosuvastatin)*, and that had not been reflected in the corresponding indent, and

ii. The indenting database uses the internationally accepted name for a drug *(Paracetamol)* whereas the error capturing database uses the American name *(Acetaminophen)* for the same drug.

Treatment of Error Categories 2,3 & 4

Since the model was to be designed based on binary classification for error occurrence prediction, it was essential to properly account for categories 2, 3 and 4.

Error Category 2 Treatment: For errors getting transposed onto multiple indents, only one of the multiple indents was chosen as the error occurrence and other was marked as category 0/

Error Category 3 Treatment: For those IPs in the indent sheet missing the drug on the given day, but an error has occurred, new indents with the error instances were created, since all other fields had known values.

Error Category 4 Treatment: A majority of the category 4 errors *(IP missing in the indent sheet, but error occurred)* were because of the actual error occurring during the discharge summary process, which happens much after the indenting process and can not be captured. Hence, these errors were dropped from the data set.

Therefore, the absorption of category 2 errors, creation of category 3 error, and the deletion of category 4 errors resulted in a dataset having 1395 erroneous indents (positive class) and ~15 lakh non erroneous indents (negative class).

Feature Engineering on Indented Data

The Need for Feature Engineering

The explanatory variables that could be used in predictive modelling very few in number, and it was essential to look at other sources of data to obtain new features, and also simultaneously look to engineer other features from the existing ones.

Some features that were desired to be engineered were:

Corrected versions of generic drug names *(standardised spellings)*
Corrected versions of medicine names *(standardised spellings)*
Medicine types *(Tablet / Injection / Capsules etc..)*
Drug allergy information
Interactions between multiple drugs
Drug substitution information
Patient contraindication to drugs *(to serve as caution/warnings)*

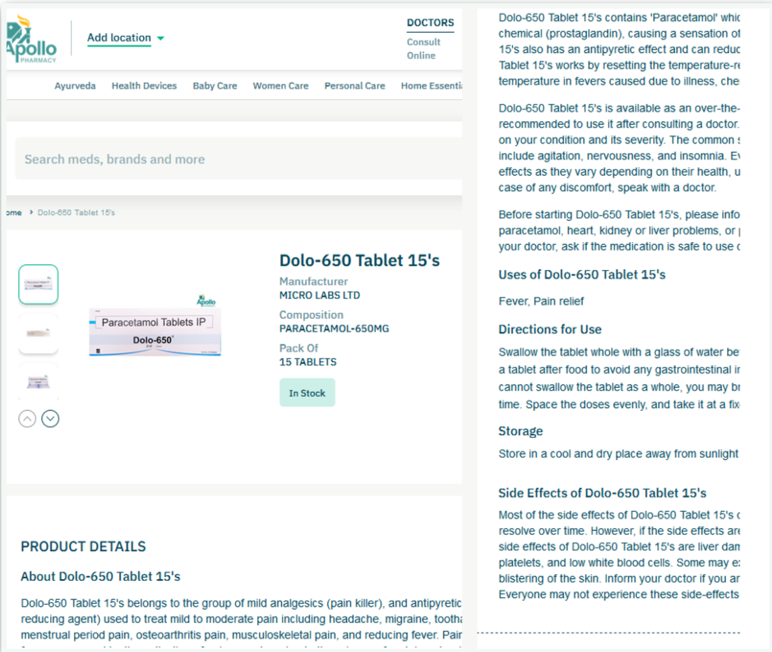
Apollo Hospitals provided their drug formulary, as well as recommended the following sources from where this information could be obtained:

rxlist.com
webmd.com
drugs.com
medscape.com

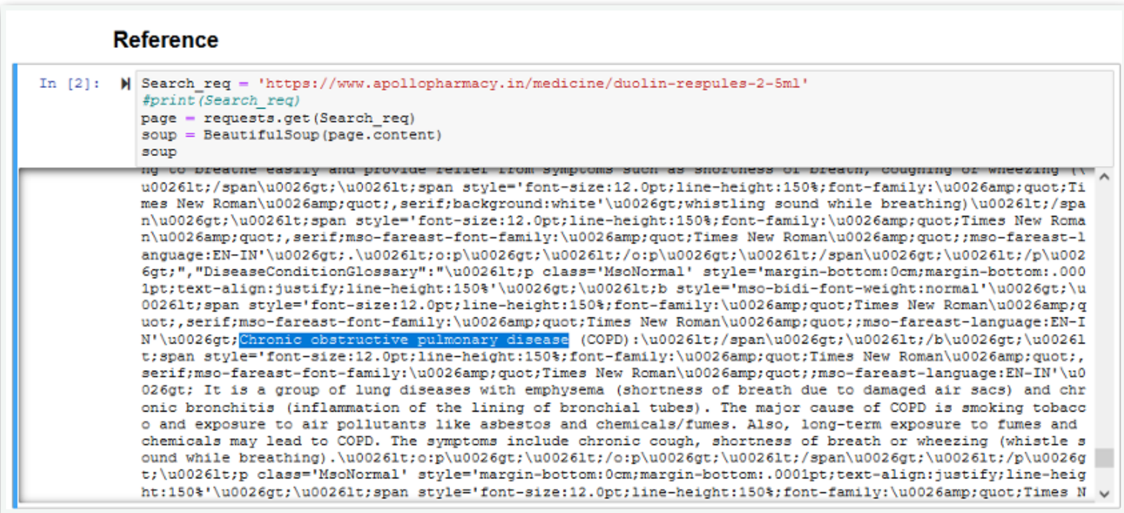
However, automatic the drug related information search was not feasible from these sites since available generic/medicine names were not standardised and had spelling errors.

To overcome this, DrugCodes specific to Apollo Hospitals that were provided, were used to identify the drug related information from Apollo Pharmacy's e-commerce website, and then use text analytics techniques to extract the relevant information, also using Selenium and BeautifulSoup packages in Python.

Product Page Navigation Through Automated Drug Search



Sample Retrieved Raw Data



Reference

```
In [2]:  Search_req = 'https://www.apollopharmacy.in/medicine/duolin-respules-2-5ml'
         #print(Search_req)
         page = requests.get(Search_req)
         soup = BeautifulSoup(page.content)
         soup
```


```
u0026lt;/span\u0026gt;\u0026lt;span style='font-size:12.0pt;line-height:150%;font-family:\u0026amp;quot;Ti
mes New Roman\u0026amp;quot;,serif;background:white'\u0026gt;whistling sound while breathing)\u0026lt;/spa
n\u0026gt;\u0026lt;span style='font-size:12.0pt;line-height:150%;font-family:\u0026amp;quot;Times New Roma
n\u0026amp;quot;,serif;mso-fareast-font-family:\u0026amp;quot;Times New Roman\u0026amp;quot;;mso-fareast-l
anguage:EN-IN'\u0026gt;.\u0026lt;o:p\u0026gt;\u0026lt;/o:p\u0026gt;\u0026lt;/span\u0026gt;\u0026lt;/p\u002
6gt;","DiseaseConditionGlossary":"\u0026lt;p class='MsoNormal' style='margin-bottom:0cm;margin-bottom:.000
1pt;text-align:justify;line-height:150%'\u0026gt;\u0026lt;b style='mso-bidi-font-weight:normal'\u0026gt;\u
0026lt;span style='font-size:12.0pt;line-height:150%;font-family:\u0026amp;quot;Times New Roman\u0026amp;q
uot;,serif;mso-fareast-font-family:\u0026amp;quot;Times New Roman\u0026amp;quot;;mso-fareast-language:EN-I
N'\u0026gt;Chronic obstructive pulmonary disease (COPD):\u0026lt;/span\u0026gt;\u0026lt;/b\u0026gt;\u00261
t;span style='font-size:12.0pt;line-height:150%;font-family:\u0026amp;quot;Times New Roman\u0026amp;quot;,
serif;mso-fareast-font-family:\u0026amp;quot;Times New Roman\u0026amp;quot;;mso-fareast-language:EN-IN'\u0
026gt; It is a group of lung diseases with emphysema (shortness of breath due to damaged air sacs) and chr
onic bronchitis (inflammation of the lining of bronchial tubes). The major cause of COPD is smoking tobacc
o and exposure to air pollutants like asbestos and chemicals/fumes. Also, long-term exposure to fumes and
chemicals may lead to COPD. The symptoms include chronic cough, shortness of breath or wheezing (whistle s
ound while breathing).\u0026lt;o:p\u0026gt;\u0026lt;/o:p\u0026gt;\u0026lt;/span\u0026gt;\u0026lt;/p\u0026g
t;\u0026lt;p class='MsoNormal' style='margin-bottom:0cm;margin-bottom:.0001pt;text-align:justify;line-heig
ht:150%'\u0026gt;\u0026lt;span style='font-size:12.0pt;line-height:150%;font-family:\u0026amp;quot;Times N
```

Information Retrieval (IR) Process For Text Mining & Text Analytics

Text mining, also known as text data mining, is the process of transforming unstructured text into a structured format to identify meaningful patterns and new insights. Text mining and text analysis identifies textual patterns and trends within unstructured data through the use of machine learning, statistics, and linguistics. By transforming the data into a more structured format through text mining and text analysis, more quantitative insights can be found through text analytics.

Information retrieval (IR) returns relevant information or documents based on a pre-defined set of queries or phrases. IR systems utilise algorithms to track user behaviours and identify relevant data. Information retrieval is commonly used in library catalogue systems and popular search engines, like Google. Some common IR sub-tasks include:

- **Tokenisation:** This is the process of breaking out long-form text into sentences and words called "tokens". These are, then, used in the models, like bag-of-words, for text clustering and document matching tasks.

- **Stemming:** This refers to the process of separating the prefixes and suffixes from words to derive the root word form and meaning. This technique improves information retrieval by reducing the size of indexing files.

The **information retrieval process** for obtaining the desired features as discussed in the previous section, is as follows:

i. All the unique DrugCodes in the indenting sheet are identified.

ii. The list of DrugCodes is used to search the website to ascertain the individual product.

iii. After navigating to the product page, the information dump is scraped on the page, using HTML XPath and the result is stored.

iv. The information dump is parsed and information blocks containing relevant information about the desired features is identified.

v. Precise feature related information is extracted by running a text mining script and is stored in the database

vi. The database is cleaned to fix redundant and missing information.

A **couple of challenges** that were faced during mining this information from Apollo Pharmacy's website were:

i. Drugs which are not over-the-counter, and those which can either accessed only by physicians / with the help of a valid prescription were unavailable for mining, and

ii. Drugs which not in stock on the e-commerce website, and certain restricted drugs also could not be extracted.

Text Analytics Techniques

In order to create the database, the following text analytics and mining techniques were implemented:

i. **Regular Expressions**, like Parsing HTML, for dimensionality reduction and information retrieval.

ii. **Topic Modelling** to identify the context behind the text block *(for example: Drug Interactions)*

iii. **Parts of Speech (POS) tagging** to extract the entities, *(like the exact name for allergies)*

iv. **Named Entity Recognition (NER)** to identify and differentiate between different entities (like Generic Name vs Branded Medicine Name)

v. **Sentiment Analysis** to filter out phrases of interest, from the generic information dump. *(Like, for example, "unsafe for pregnant women" / "avoid ______ when" / "reconsider dosage")*

Sample Images of the Database Created from Apollo Pharmacy Website using Text Analytics

Drug Information

DRUGCODE	MEDICINENAME	MedicineType	Corrected_GenericNames
THY0005	THYROX 50MCG TAB	TABLETS	THYROXINE SODIUM
PUL0108	PULMOCLEAR SYP 100ML	SYRUP	ACEBROPHYLLINE+MENTHOL+TERBUT/
PHE0027	PHENERGAN 100ML SYP	SYRUP	PROMETHAZINE
MYO0002	MYOSPAZ TAB	TABLETS	CHLORZOXAZONE+PARACETAMOL
TEL0062	TELLZY 40MG TAB	TABLETS	TELMISARTAN
LEV0059	LEVIPIL 5ML INJ	INJECTION	LEVETIRACETAM
GLU0022	GLUMET TAB	TABLETS	METFORMIN
TEL0017	TELMA 40MG TAB 15'S	TABLETS	TELMISARTAN
BET0010	BETALOC 25MG TAB	TABLETS	METOPROLOL
ABF0001	ABFLO 100MG CAP	CAPSULES	ACEBROPHYLLINE
GLY0024	GLYCOMET 500MG TAB	TABLETS	METFORMIN
TAX0010	TAXIM-O 200 MG TAB	TABLETS	CEFIXIME
GLY0028	GLYCOMET SR 500MG TA	TABLETS	METFORMIN

Drug Interaction

Drug1	Drug2	DI_Code
ACECLOFENAC+DROTAVERINE	IBUPROFEN	0
ACECLOFENAC+DROTAVERINE	DICLOFENAC	0
ACECLOFENAC+DROTAVERINE	NAPROXEN	0
ACECLOFENAC+DROTAVERINE	KETOROLAC	0
ACECLOFENAC+PARACETAMOL+THIOCOLCHICOSIDE	IBUPROFEN	0
ACECLOFENAC+PARACETAMOL+THIOCOLCHICOSIDE	DICLOFENAC	0
ACECLOFENAC+PARACETAMOL+THIOCOLCHICOSIDE	NAPROXEN	0

Patient Contraindications for Drugs (0- No Issues, 1- Strictly Avoid, 2 - Reconsider Dosage)

GenericName	Pregnancy_flag	Kidney_flag	Liver_flag	Breastfeeding_flag
THYROXINE SODIUM	2	2	2	2
ACEBROPHYLLINE+MENTHOL+TERBUTALINI	0	0	0	0
PROMETHAZINE	2	2	2	1
CHLORZOXAZONE+PARACETAMOL	2	2	1	2
TELMISARTAN	2	2	1	2
LEVETIRACETAM	1	2	2	1
METFORMIN	1	1	2	1
METOPROLOL	1	2	2	2
ACEBROPHYLLINE	1	2	2	2
CEFIXIME	1	2	2	2

Drug Allergies

Generic name	Allergy_flag
ACECLOFENAC	ibuprofen, diclofenac, naproxen, ketorolac
AMIKACIN	antibiotics related to aminoglycosides, sulphites
AMIODARONE	iodine, amiodarone hydrochloride
ASPIRIN+CLOPIDOGREL	clopidogrel, aspirin, salicylates, anti-inflammatories (pain killers)
ASPIRIN+ROSUVASTATIN	aspirin and rosuvastatin
ATORVASTATIN+EZETIMIBE	sugar
AZATHIOPRINE	immunosuppressant medicines or Azathioprine
AZILSARTAN	azilsartan medoxomil
AZILSARTAN MEDOXOMIL	azilsartan medoxomil

Final Exploratory Data Analysis

Objectives of Final Exploratory Data Analysis

Exploratory data analysis was carried out on the indented data that was cleaned, mapped, and feature engineered, to identify top contributors under each category, that could help further reduce data dimensionality, before proceeding with modelling.

Part I: Exploratory Data Analysis on 15 Lakh indents

Fields Used: IndentCount, ErrorCount, DrugCode, WardName, GenericName, MedicineType

Period: 2019

Key EDA insights from Indents

i. **Drugs:**

  a. There are **12 drugs** *(out of 8683 drugs)* which are prescribed **more than 10,000 times** in the year, and they account for **~20% of all the indents**.

  b. There are **48 drugs** *(out of 8683 drugs)* causing more than 5 errors in the year, and these account for **~50% of all the errors** that occurred.

  c. **8 drugs** are **critical in terms of error proneness,** with high number of errors in the year, as well as causing errors more than 10% of the times they were prescribed. Out of these 8, **<u>P120001 & GLU0022 have an extremely high error percentage of 94.5% and 80.5% respectively.</u>**

ii. **Wards:**

  a. **5 wards** *(out of 39 wards)* have more than 1 Lakh prescriptions in the year. These accounted for **~53% of all prescriptions** in the year.

  b. **More than 100 errors per ward, happened in just 6 wards** *(out of 39 wards)* in the year. These accounted for **~70% of all the errors** in the year.

iii. **Generic Names:**

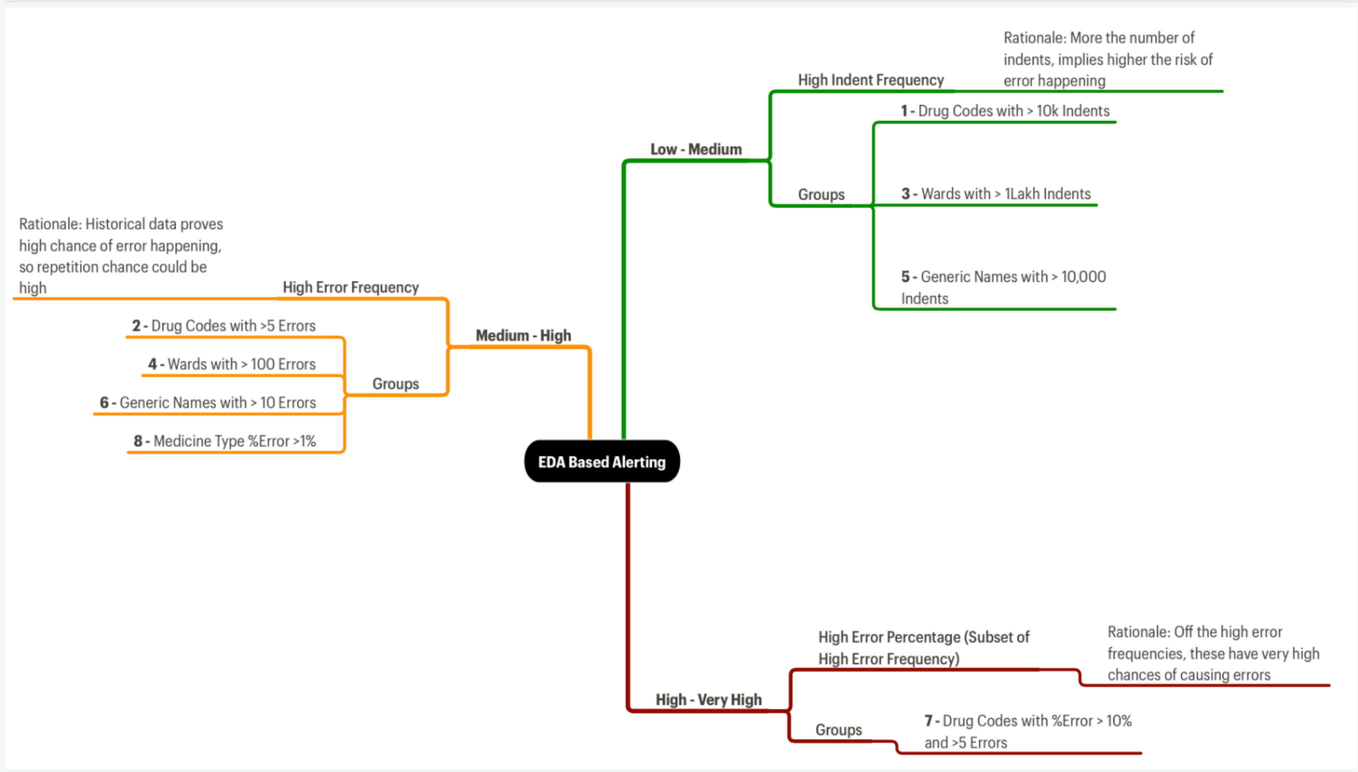
  a. **21 Generic Names,** *(out of 2375 Generic Names)* were prescribed more than 10,000 times in the year, accounting for **~38% of all prescriptions**.

  b. **25 Generic Names,** *(out of 2375 Generic Names)* caused more than 10 errors in the year, accounting for **~55% of all errors.**

iv. **Medicine Type:**

  a. The medicine type “Suspension” *(out of 28 Medicine Types)* had an error % of 1.484%, whereas the median error % for the rest of medicine types is 0.054%

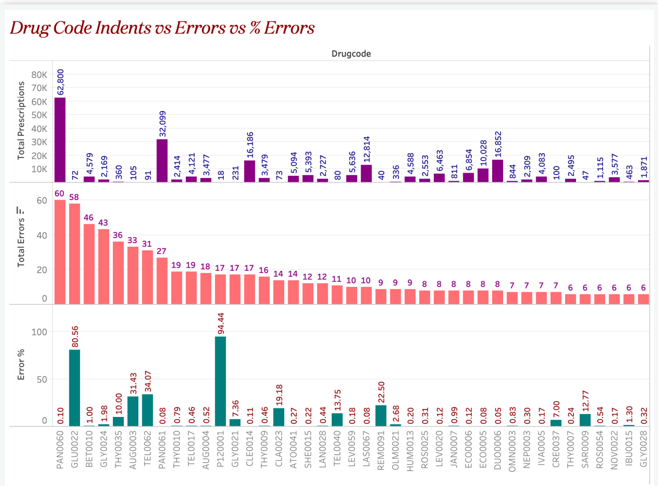
The insights from the Indenting EDA have been used to form groups of top percentile values under each field to form three categories of criticality, “low-medium”, “medium-high”, and “high-very high”, which could be used to group the overall indents, while building predicting models, as well as for altering systems, as shown below:

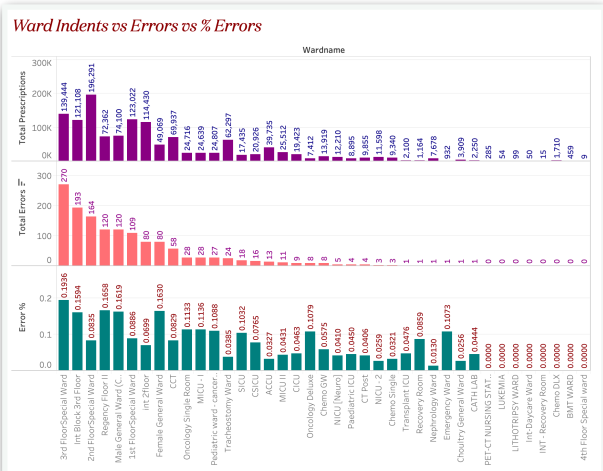


Sample Graphs from the Final EDA for 15 Lakh Indents

Indents vs Errors vs Error % For Drugs & Wards.

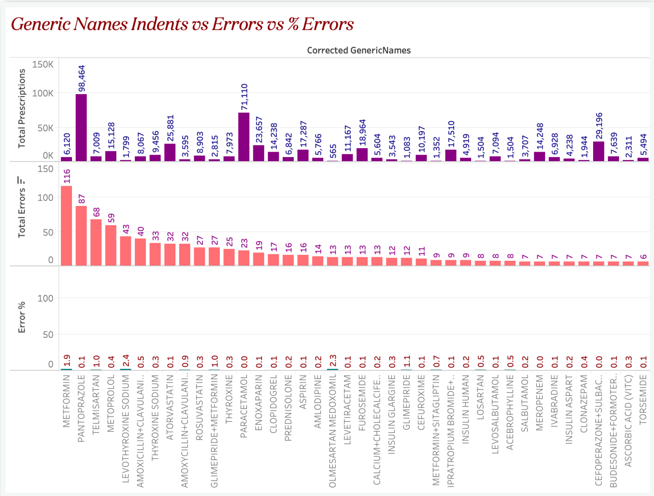
(Sorted descending by Error Count)

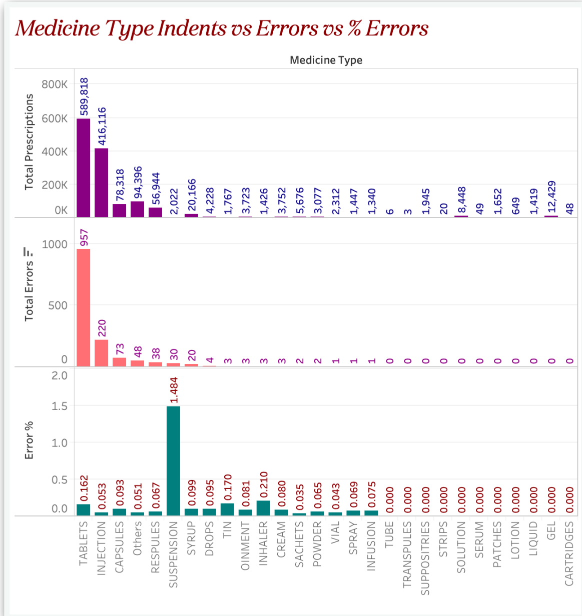




Indents vs Errors vs Error % For Generic Names & Medicine Types

(Sorted descending by Error Count)





Part II: Exploratory Data Analysis on 1395 Errors

Since the scraped data from Apollo Hospitals' website helped in refining the drug names, and agglomerating dosage of drugs and their strengths, as a part of feature engineering, EDA was also carried out once again on the final error sheet for the one Group-A hospital unit.

Fields Used: ErrorSubType, ErrorCause, DrugCategory, Specialty

Period: 2019

Key EDA Insights From Errors

i. **Error Sub Type:**

a. ~78% of all errors happen because of 5 subtypes. *(with >50 errors per subtype)*

b. ~80% of all prescription errors can be ascribed to 5 sub types. *(with > 50 prescription errors per subtype)*

c. ~64% of all transcription errors can be ascribed to 5 sub types. *(with > 10 transcription errors per subtype)*

d. ~78% of all administration errors can be ascribed to 4 sub types. *(with > 2 administration errors per subtype)*

ii. **Error Cause:**

a. ~83% of all errors happen because of 4 causes. *(with > 100 total errors per cause)*

b. ~87% of all prescription errors can be ascribed to 5 causes. *(with > 100 prescription errors per cause)*

c. ~74% of all transcription errors can be ascribed to 5 causes. *(with > 20 transcription errors per cause)*

d. ~70% of all administration errors can be ascribed to 4 causes. *(with > 3 administration errors per cause)*

iii. **Specialty:**

a. ~74% of all errors happen because of 7 specialties. *(with > 80 total errors per specialty)*

iv. **Drug Category:**

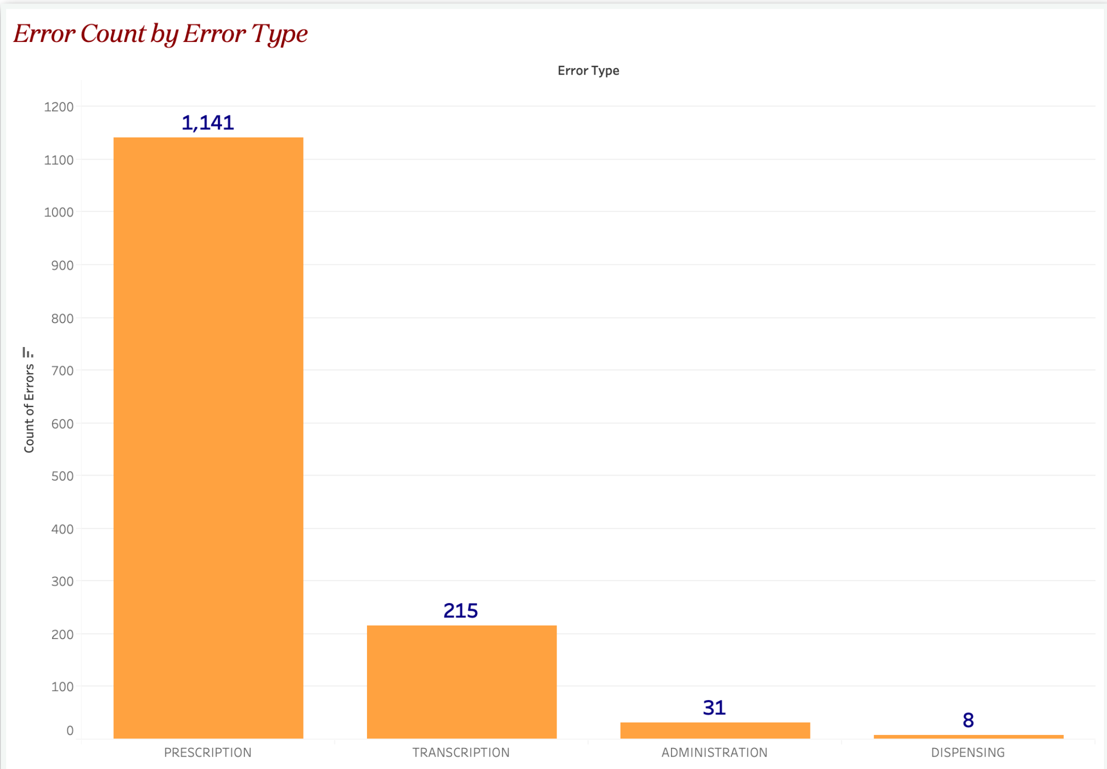
a. ~40% of all errors happen because of 9 causes. *(with > 25 total errors per cause)*

The insights from the error sheet after cleaning and mapping with indents, are similar to the insights derived during the initial EDA conducted on the 6 marquee hospitals identified by Apollo Hospitals.

Sample Graphs from the Final EDA for 1395 Errors

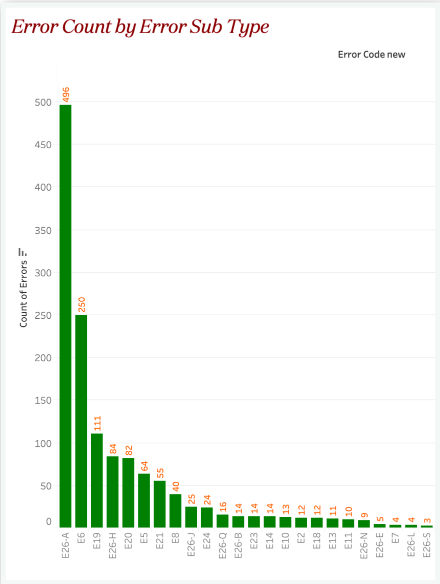
Error Count vs Error Type

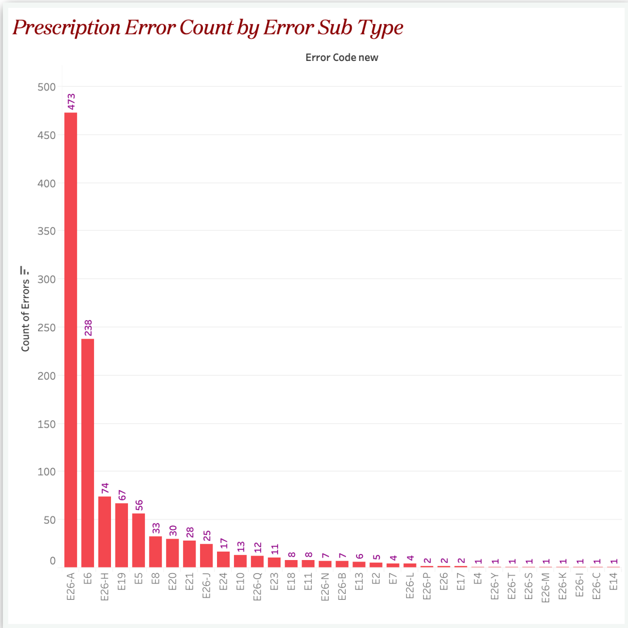
(Sorted descending by Error Count)



Error Count vs Error Sub Type (Total vs Prescription)

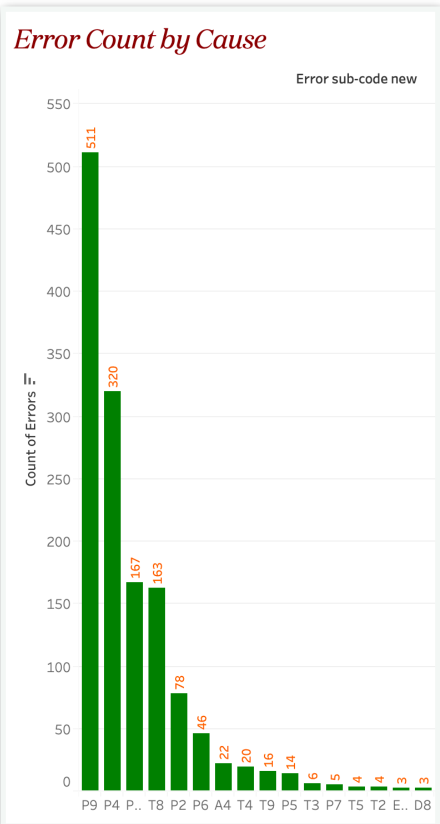
(Sorted descending by Error Count)

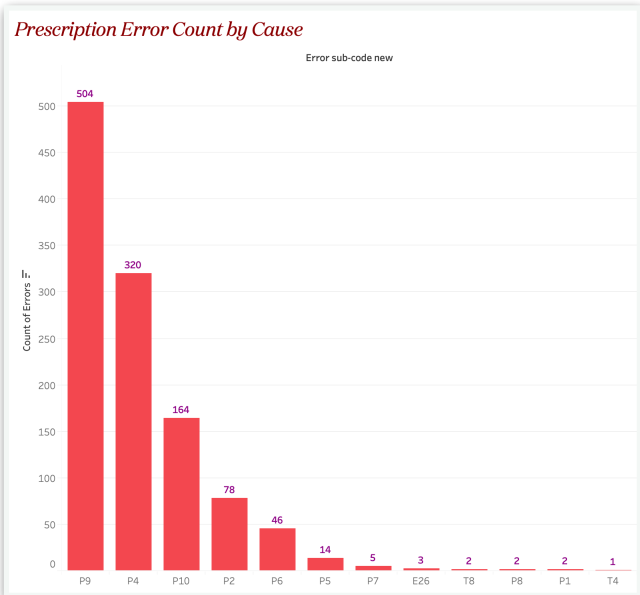




Error Count vs Error Cause (Total vs Prescription)

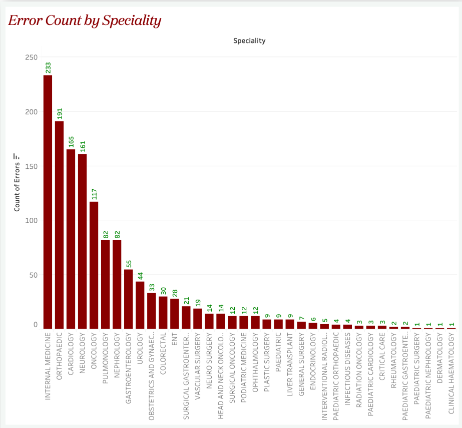
(Sorted descending by Error Count)

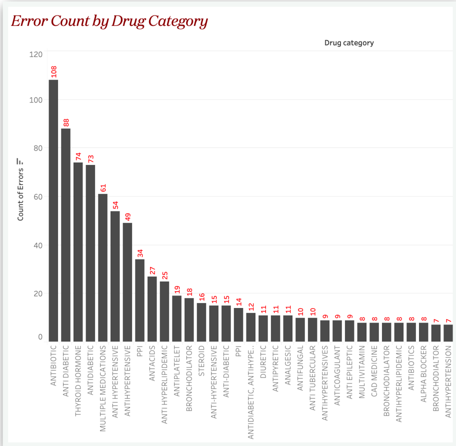




Error Count vs Specialty & Drug Category

(Sorted descending by Error Count)





Dimensionality Reduction & Classification - Supervised Learning Models

Preparation:

After cleaning & mapping the errors with the indents, feature engineering new variables, and performing a final EDA, the data was in format, ready for classification modelling, with 1395 positive class (errors) and ~15 lakh negative class (non error) records.

Hypothesis Testing: The zero class and one class data had 2 features which were both categorical, high in number of unique categories and extremely skewed, Chi-Squared test for Independence did not yield interpretable results, leading to multiple !Div/0 Errors.

Resampling: The data was severely imbalanced with the positive class representing ~0.1% of the data. Thus, multiple samples were created with different resampling strategies.

Dimensionality Reduction: Since there were only two predictors (drug & ward), describing thousands of resampled records, there was no need to further reduce the number of fields.

Clustering: K-Modes clustering was used for multiple values of K, but cluster separation achieved was inefficient, leading to the number of clusters being too few or too many.

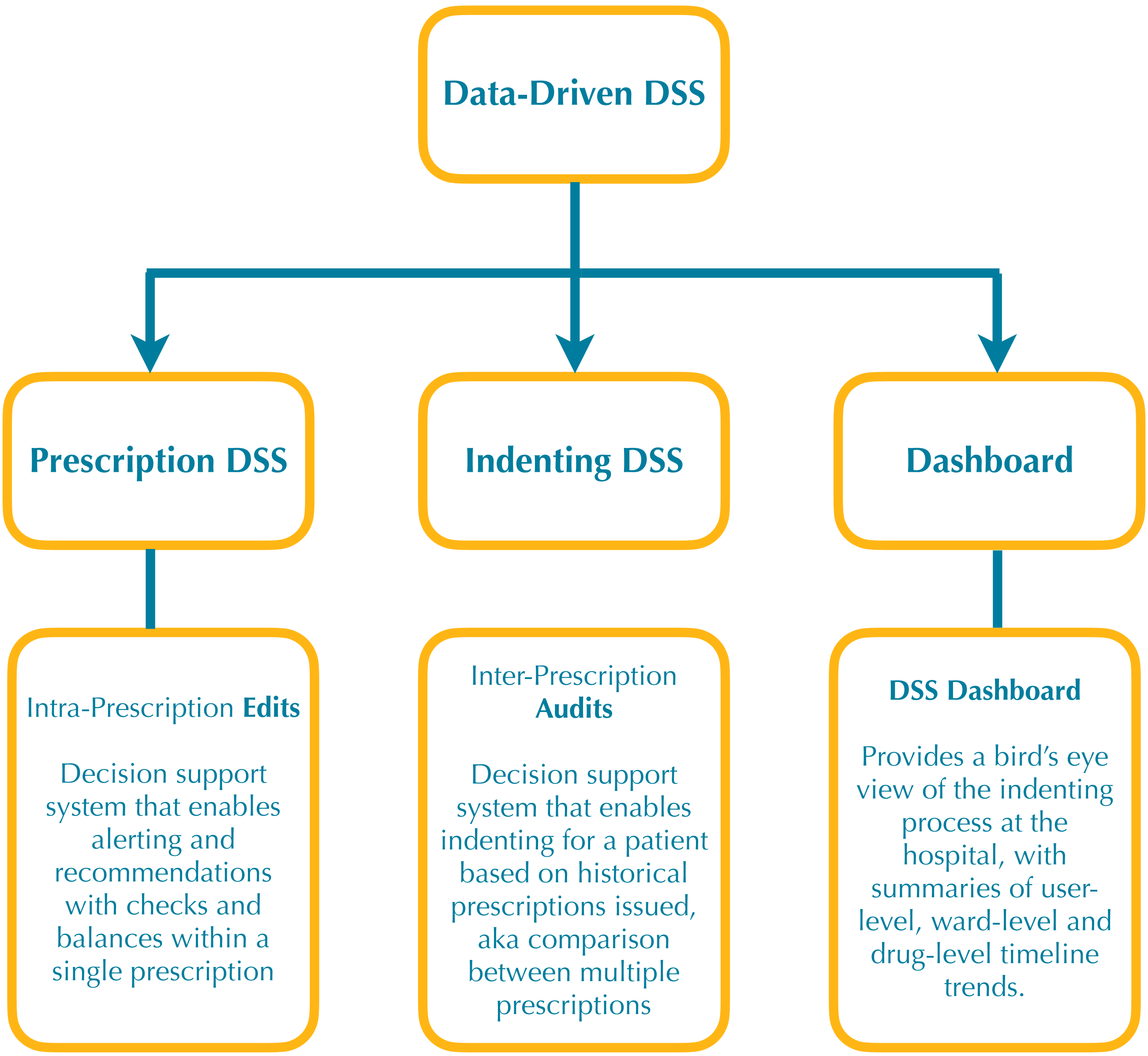
Classification: Classifications algorithms such as Naive Bayes, Logistic Regression, Decision Trees and SVM were experimented with, but the none of the results were comprehensible, with little to no class separation possible, owing to the fact that there were just too few features to explain the variation in the outcome variable *(error occurrence)*

Way Ahead: After deliberation with Apollo Hospitals, it was decided that prediction modelling was infeasible, given the lack of richness in the data and instead, **a rule-based decision support system** would be built, incorporating both insights from the different stages of EDA, and Apriori algorithm to address the four main critical error factors from FMEA, among other error causes.

Building A Data-Driven Decision Support System

Structure of the DSS

A data-driven decision support system (DSS) was built with the following structure and base functionalities, as shown below. This would facilitate Apollo Hospitals to be more informed about potential medication errors in the future, and take actions that directly help prevent adverse patient eventualities arising that could arise from medication errors.



The above three systems; functionalities are powered by business rules, which allow the DSS to intimate key process owners at Apollo Hospitals about potential risks in medication.

DSS Functions & the Business Rules

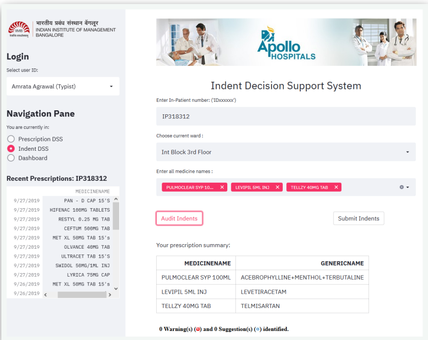
Prescription DSS - The Basic Functions

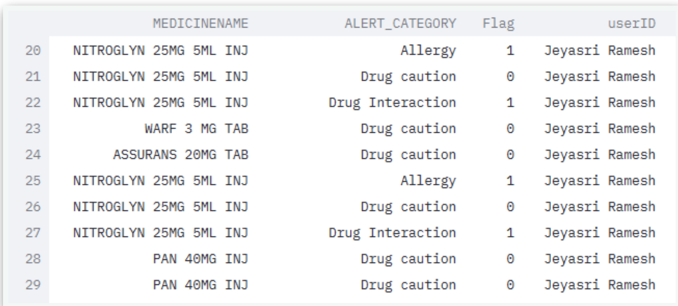
The basic functional flow of the DSS application allows:

iii. Individual users to login under their specific IDs and indent their respective prescriptions.
iv. Users to first enter the IP number under which the prescription is to be loaded.
v. Users to a enter list of drugs, either by typing the actual brand name *(for nurses/typists with doctor's prescription in hand)* or by searching for brand names given a generic name *(for doctors while prescribing)*
vi. Users to view a list of alerts - suggestions *(flag=0)* if there are no red flags, and warnings *(flag=1)*, if there are red flags under each drug indented.
vii. Users to remove alert-inducing drugs, and submit the cleared indent, or to overrule/ignore the alerts and submit anyway.
viii. Users to view a prescription summary with all the drugs selected under each indent.

<u>Every submitted indent will further be categorised, with an override *(alerts ignored)* or not *(no alerts)*.</u>

Prescription DSS - Basic Functions - Frontend & Backend





	MEDICINENAME	ALERT_CATEGORY	Flag	userID
20	NITROGLYN 25MG 5ML INJ	Allergy	1	Jeyasri Ramesh
21	NITROGLYN 25MG 5ML INJ	Drug caution	0	Jeyasri Ramesh
22	NITROGLYN 25MG 5ML INJ	Drug Interaction	1	Jeyasri Ramesh
23	WARF 3 MG TAB	Drug caution	0	Jeyasri Ramesh
24	ASSURANS 20MG TAB	Drug caution	0	Jeyasri Ramesh
25	NITROGLYN 25MG 5ML INJ	Allergy	1	Jeyasri Ramesh
26	NITROGLYN 25MG 5ML INJ	Drug caution	0	Jeyasri Ramesh
27	NITROGLYN 25MG 5ML INJ	Drug Interaction	1	Jeyasri Ramesh
28	PAN 40MG INJ	Drug caution	0	Jeyasri Ramesh
29	PAN 40MG INJ	Drug caution	0	Jeyasri Ramesh

Generating the Alerts - Business Rules

There are principal conditions to be met by each drug either in isolation, or in combination with other drugs, to be cleared as potentially error-free. These conditions are driven by two sets of knowledge-bases that have been created just for this purpose.

The first knowledge-base is powered by Apriori algorithm, and the second knowledge-base is powered by drug-level anomalies which can be detected through the database formed from scraping Apollo Pharmacy's website.

<u>Apriori is an unsupervised learning algorithm for item set mining and generating association rules. It is most commonly used in market basket analysis, and in this scenario, has been implemented for understanding the association between various drugs in their capacity to cause potential medication errors.</u>
<u>Apriori has been used to compute "**support**" between two given drugs, which measures the number of times the two given drugs are prescribed together, and **"confidence"**, which is</u>

measures the conditional probability of the first drug being prescribed, given that the other is also prescribed.

Apriori algorithm accounts for the first two business rules, as listed below:

Business Rule #1 : Apriori - ALL

Use - The first business rule is to used primarily to provide the users with suggestions on probable drugs that they may have missed, based on set thresholds for support and confidence.
Benefit - This will help users avoid common mistakes such as failing to add one or more drugs that are routinely added in combination, which can in turn save time, along with minimising error chances.

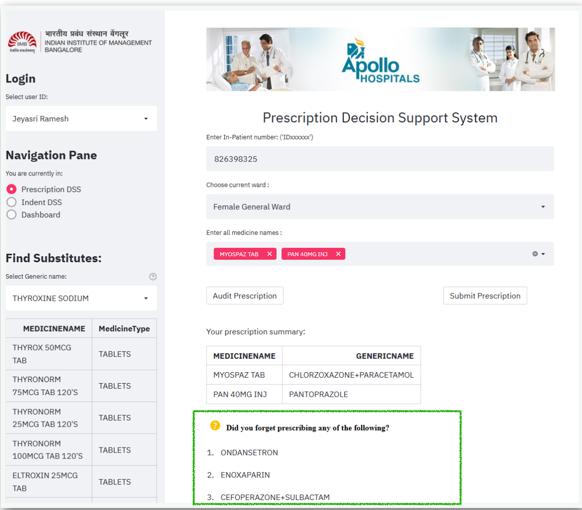
Business Rule #2 : Apriori - EDA Embedded

Use - This rule also follows Apriori, and comes into effect when the previous rule fails to identify specific drugs combinations. This rule uses insights (***top error contributing drugs / wards, etc.***) from various levels of exploratory data analysis (EDA) to identify drug combinations localised to trends within those specific fields, like for example, the indenting for a specific location in the hospital considers only trends from that ward, and checks if the top error contributing drugs in that ward are present in this combination or not.

This business rule is intended to account for **recency** and **frequency** of drugs being prescribed. Since dates and times haven't been effectively captured for the previous indents, only frequency information from wards are incorporated, however, in the future, with future indents, both recency and frequency can be taken into account.

Benefit - This is unique to every hospital and is dependent on historical indents that have been provided. While it can provide insights from only those drugs previously indented, it also enables the hospital to make the DSS "learn" from the previous indents, and to provide more refined alerts in the future.

Apriori Algorithm-Based Business Rules - Frontend ***(in green)*** **& Backend** ***(showing support values)***



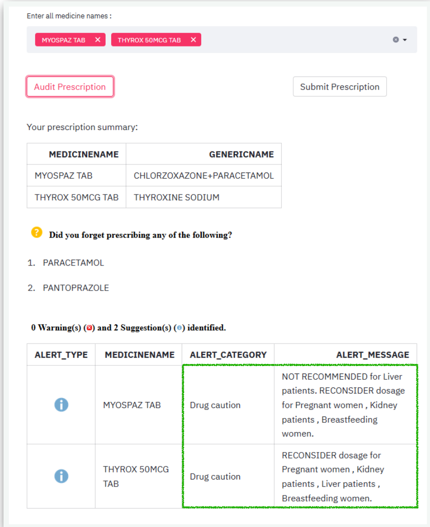
	A	B
1	support	itemsets
2	0.6139474	PANTOPRAZOLE, PARACETAMOL
3	0.436114086	PANTOPRAZOLE, ONDANSETRON
4	0.350668412	ONDANSETRON, PARACETAMOL
5	0.321750169	PANTOPRAZOLE, ONDANSETRON, PARACETAMOL
6	0.256813043	PANTOPRAZOLE, CEFOPERAZONE+SULBACTAM
7	0.225276687	CEFOPERAZONE+SULBACTAM, PARACETAMOL
8	0.212662144	PANTOPRAZOLE, LIDOCAINE
9	0.208695307	PANTOPRAZOLE, BENZYL ALCOHOL+DICLOFENAC SODIUM
10	0.20472847	PANTOPRAZOLE, ENOXAPARIN
11	0.203419414	PANTOPRAZOLE, CEFOPERAZONE+SULBACTAM, PARACETAMOL
12	0.184576937	LIDOCAINE, PARACETAMOL
13	0.17743663	BENZYL ALCOHOL+DICLOFENAC SODIUM, PARACETAMOL
14	0.167559205	ENOXAPARIN, PARACETAMOL

Business Rule #3 : Patient - Drug Contraindications

Use - This rule is derived from the knowledge base generated from scraping Apollo Pharmacy's data. The information scraped from the website was used to create three kinds of contraindication flags based on the description of the individual medicine name/drug. Three contraindication flags are **0 - "No Issues", 1 - "Strictly Avoid," and 2 - "Reconsider Dosage"**. In case the contraindications involve an allergy, the flag raised is a warning.

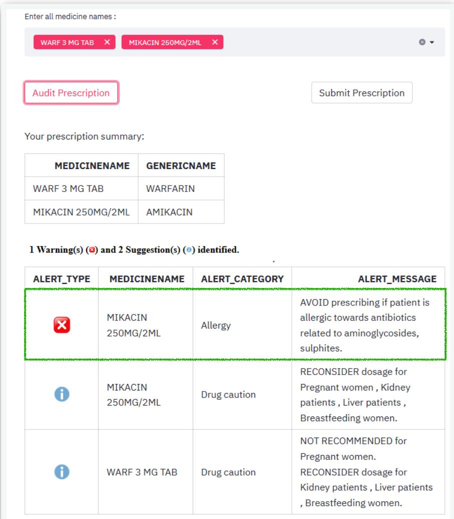
<u>**These phrases have been extracted with the help of Natural Language Processing (NLP) techniques such as Information Retrieval, POS tagging, Named Entity Recognition (NER) and Sentiment Analysis.**</u>

Benefit - Prescribing a drug when it is not supposed to given to a patient due to underlying preconditions could be dangerous. There are two types of contraindications, absolute and relative, involving either a single substance or a combination of substances that could cause harm relative to each other.
This rule assumes that the personnel prescribing the drug is aware of the preconditions, and in the future, if Apollo Hospitals wish to integrate the DSS with patient level information too, the system can then be scaled up to automatically detect drug contraindications.



Drug Contraindications - Frontend *(in green)* **& Backend** *(showing flags for contraindications)*

	A	B	C	D	E
1	GenericName	Pregnancy_flag	Kidney_flag	Liver_flag	Breastfeeding_flag
2	THYROXINE SODIUM	2	2	2	2
3	ACEBROPHYLLINE+MENTHOL+TERBUTALIN	0	0	0	0
4	PROMETHAZINE	2	2	2	1
5	CHLORZOXAZONE+PARACETAMOL	2	2	1	2
6	TELMISARTAN	2	2	1	2
7	LEVETIRACETAM	1	2	2	1
8	METFORMIN	1	1	2	1
9	METOPROLOL	1	2	2	2
10	ACEBROPHYLLINE	1	2	2	2



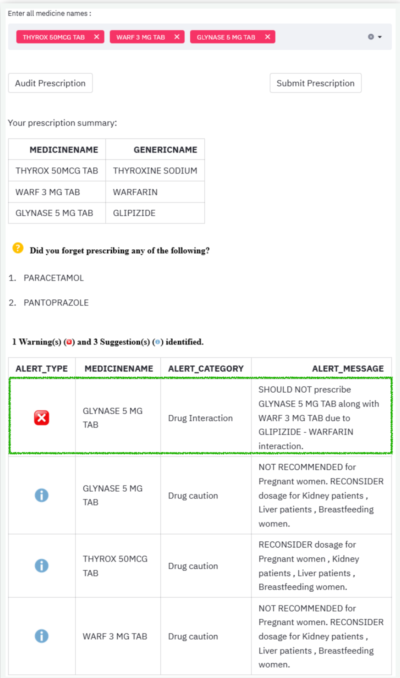
Drug Allergies - Frontend *(in green)* **& Backend** *(showing flags for allergies)*

	A	B
1	Generic name	Allergy_flag
2	ACECLOFENAC	ibuprofen, diclofenac, naproxen, ketorolac
3	AMIKACIN	antibiotics related to aminoglycosides, sulphites
4	AMIODARONE	iodine, amiodarone hydrochloride
5	ASPIRIN+CLOPIDOGREL	clopidogrel, aspirin, salicylates, anti-inflammatories (pain killers)
6	ASPIRIN+ROSUVASTATIN	aspirin and rosuvastatin
7	ATORVASTATIN+EZETIMIBE	sugar
8	AZATHIOPRINE	immunosuppressant medicines or Azathioprine
9	AZILSARTAN	azilsartan medoxomil
10	AZILSARTAN MEDOXOMIL	azilsartan medoxomil
11	BECLOMETASONE+CLOTRIMAZOLE+NEOMYCIN	steroid medicines and antibiotics
12	BETAMETHASONE+ZINC SULFATE	steroid medicine

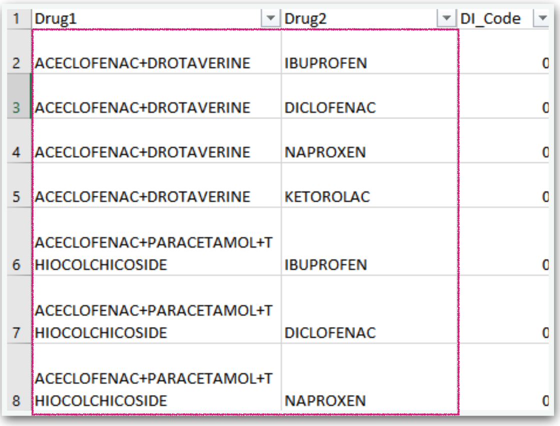
Business Rule #4 : Drug Interactions

Use - Similar to contraindications, the Apollo Pharmacy website was scraped to also identify patterns in the data which describe possible harmful interactions between drugs. This has been incorporated into a business rule that disallows indenting personnel from mistakenly issuing two or more drugs that when combined can harm patients.

Benefit - The system automatically raises warnings *(red flags)* to drugs whose interaction that can cause harm. The user is then provided the option to remove the error inducing drug and refine the prescription, or override this warning and submit the prescription, which, as mentioned before, is captured for future monitoring and control.



Drug Interactions - Frontend *(in green)* & Backend
(showing flags for combinations of drugs that may interact)



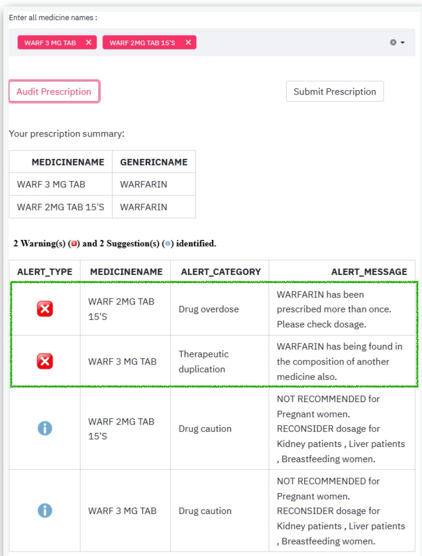
1	Drug1	Drug2	DI_Code
2	ACECLOFENAC+DROTAVERINE	IBUPROFEN	0
3	ACECLOFENAC+DROTAVERINE	DICLOFENAC	0
4	ACECLOFENAC+DROTAVERINE	NAPROXEN	0
5	ACECLOFENAC+DROTAVERINE	KETOROLAC	0
6	ACECLOFENAC+PARACETAMOL+THIOCOLCHICOSIDE	IBUPROFEN	0
7	ACECLOFENAC+PARACETAMOL+THIOCOLCHICOSIDE	DICLOFENAC	0
8	ACECLOFENAC+PARACETAMOL+THIOCOLCHICOSIDE	NAPROXEN	0

Business Rule #5 : Inadvertent Drug Duplications

Use - Not unlike contraindications or drug interactions, it is not uncommon for a prescription to erroneously have drug duplications *(as seen in error indents)*.To avoid this, the DSS shall raise flags whenever two or more drugs having the same therapeutic constituents is entered into one prescription.

Benefit - By accounting for drug duplications in indents, a possible drug overdose can be avoided.

Drug Duplications - Frontend
(in green)



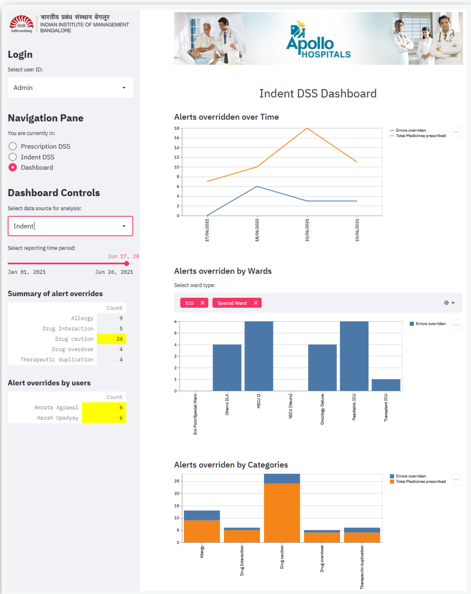
Prescription DSS vs Indenting DSS

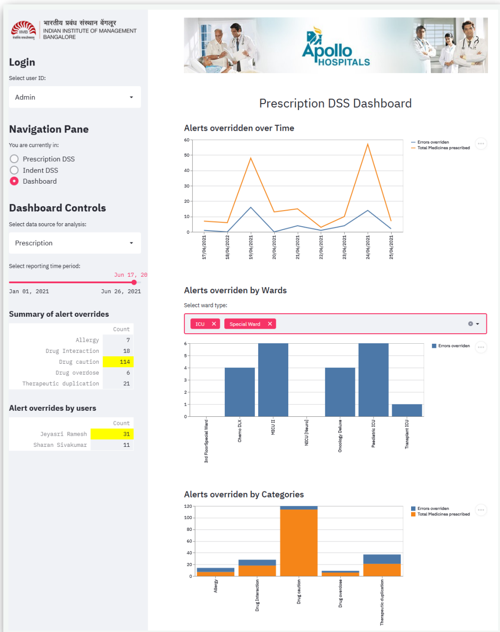
The business rules for the Indenting DSS are the same as that of the Prescription DSS, except that the Indenting DSS takes an extra business rule, which checks the relationship of the current indent being raised for a patient vis à vis all the previous indents that have been raised for this patient during this admission tenure. Prescription DSS has been tailor made for the use of doctors and the Indenting DSS for nurses/typists, but can be used interchangeably.

By evaluating the relevancy of these business rules, it can be seen that the bottom three rules purely involve drug dynamics and is common to any hospital unit. But it is these three rules in combination with the Apriori-based first two rules that make this system unique for every hospital.

Monitoring & Reporting Alerts with the DSS Dashboard

A dashboard was created to enable the superuser *(quality team at Apollo)* to monitor the performance of the personnel using the DSS, as shown below:





The dashboard gives user-wise alert overrides, trends of overridden alerts over time, for selected wards *(SICU, Oncology..)*, **and across the different categories of alerts** *(Allergy, Therapeutic Duplication..)* **for both indenting & prescription DSS.**

MEDICINENAME	ALERT_CATEGORY	Flag	userID	page	date	time	ward
MYOSPAZ TAB	Drug caution	1	Harsh Upadyay	Indent	22/06/2021	21:14:00	PET-CT NURSING STATION
LEVIPIL 5ML INJ	Drug overdose	0	Harsh Upadyay	Indent	22/06/2021	21:14:00	PET-CT NURSING STATION
PHENERGAN 100ML SYP	Drug caution	0	Harsh Upadyay	Indent	22/06/2021	21:14:00	PET-CT NURSING STATION
ABFLO 100MG CAP	Drug caution	0	Harsh Upadyay	Indent	22/06/2021	21:14:00	PET-CT NURSING STATION
MYOSPAZ TAB	Drug caution	0	Harsh Upadyay	Indent	22/06/2021	21:14:36	PET-CT NURSING STATION
LEVIPIL 5ML INJ	Therapeutic duplic	0	Harsh Upadyay	Indent	22/06/2021	21:14:36	PET-CT NURSING STATION
PHENERGAN 100ML SYP	Therapeutic duplic	0	Harsh Upadyay	Indent	22/06/2021	21:14:36	PET-CT NURSING STATION
ABFLO 100MG CAP	Drug caution	0	Harsh Upadyay	Indent	22/06/2021	21:14:36	Transplant ICU
MYOSPAZ TAB	Drug caution	0	Harsh Upadyay	Indent	22/06/2021	21:19:36	Transplant ICU
LEVIPIL 5ML INJ	Drug caution	0	Harsh Upadyay	Indent	22/06/2021	21:19:36	Transplant ICU
PHENERGAN 100ML SYP	Drug overdose	0	Harsh Upadyay	Indent	22/06/2021	21:19:36	Transplant ICU
ABFLO 100MG CAP	Drug caution	0	Harsh Upadyay	Indent	22/06/2021	21:19:36	Transplant ICU
MYOSPAZ TAB	Drug caution	0	Harsh Upadyay	Indent	22/06/2021	21:20:17	Transplant ICU
LEVIPIL 5ML INJ	Drug overdose	0	Harsh Upadyay	Indent	22/06/2021	21:20:17	Transplant ICU
PHENERGAN 100ML SYP	Drug caution	0	Harsh Upadyay	Indent	22/06/2021	21:20:17	Transplant ICU
ABFLO 100MG CAP	Drug caution	0	Harsh Upadyay	Indent	22/06/2021	21:20:17	Transplant ICU
WARF 3 MG TAB	Drug caution	0	Jeyasri Ramesh	Prescription	22/06/2021	21:22:34	Transplant ICU
ASSURANS 20MG TAB	Drug caution	0	Jeyasri Ramesh	Prescription	22/06/2021	21:22:34	INT - Recovery Room

The dashboard also gives an overall summary of alert overrides summed across each category of alerts *(Contraindication, Interaction..)*. **This dashboard is powered by a log of all the alerts and the indents at the backend, as shown above.**

Creating a Knowledge Base With Natural Language Processing for the Business Rules

Alerting System from Scraped Website Data

Here's a sample of the raw data dump scraped from the Apollo Pharmacy website:


```
<script class="structured-data-list" type="application/ld+json">{"@context":"https://schema.org/","@type":"Drug","name":"
Duolin Respules 2.5Ml","mainEntityOfPage":"https://www.apollo247.com/medicine/duolin-respules-2-5ml","image":"https://new
assets.apollo247.com/pub/media/thumbnail/apollo24logo.png","activeIngredient":"IPRATROPIUM BROMIDE+ LEVOSALBUTAMOL-undefi
nedMCG+ MG","alcoholWarning":"","availableStrength":{"@context":"http://schema.org/","@type":"DrugStrength","activeIngred
ient":"IPRATROPIUM BROMIDE+ LEVOSALBUTAMOL-undefinedMCG+ MG"},"breastfeedingWarning":"","pregnancyWarning":"","clinicalPh
armacology":"","dosageForm":"RESPULES","drugUnit":"MCG+ MG","foodWarning":"","isAvailableGenerically":"True","legalStatu
s":"country: india, status: Approved","overdosage":"","manufacturer":{"@type":"Organization","legalName":"CIPLA LTD"},"me
chanismOfAction":"","nonProprietaryName":"Duolin Respules 2.5Ml","isProprietary":true,"prescriptionStatus":"Available by
prescription","url":"https://www.apollopharmacy.in/medicine/duolin-respules-2-5ml"}</script>
```


On applying text analytics / mining techniques as mentioned before, the raw data dump is converted to a database as shown below.

	DRUGCODE	MEDICINENAME	PRODURL	GENERICNAME	SUBSTITUTES	DRUGTYPE	SideEffects	Directions	Allergy_flag	DrugInteraction_flag	Pregnancy	Pregnancy_flag	Breastfeeding	Liver	Kidney
1866	ACA0003	ACAMPROL TABLET	https://www.apollopharmacy.in/otc/acamprol-tablet	ACAMPROSATE-333MG	['ACAMPTAS 333MG TABLET']	TABLETS	medicines, Acamprol Tablet 6's can also cause side-effects, although not everybody experiences them. abdominal pain, diarrhea,	known to be allergic to Acamprol Tablet 6's or any other medicines, please inform your doctor. Before taking Acamprol			Acamprol Tablet 6's should be used during pregnancy only when clearly needed. Discuss the		unknown if this Acamprol Tablet 6's passes into breast milk. Inform your doctor before	Tablet 6's is not recommended in patients with a history of kidney diseases because serious side	with liver damage can safely take the Acamprol Tablet 6's because Acamprol Tablet 6's is not metaboli
604	GLU0003	GLUCOBAY 25MG TABLET	https://www.apollopharmacy.in/medicine/glucobay-25mg-tablet	ACARBOSE-25MG	['GLUCAR 25MG TABLET']	TABLETS	Like all medicines, Glucobay 25 Tablet 10's can cause some common side effects such as flatulence (gas), stomach pain or diarrhoea.	If you have severe kidney or liver disorders, ulcerative colitis or Crohnâ€™s disease (conditions causing			Glucobay 25 Tablet 10's is a Category B pregnancy drug and is generally not recomme		Avoid breastfeeding while taking Glucobay 25 Tablet 10's as it may be excreted in breast	Take Glucobay 25 Tablet 10's with caution, especially if you have a history of Kidney diseases/	Take Glucobay 25 Tablet 10's with caution, especially if you have a history of Liver diseases/

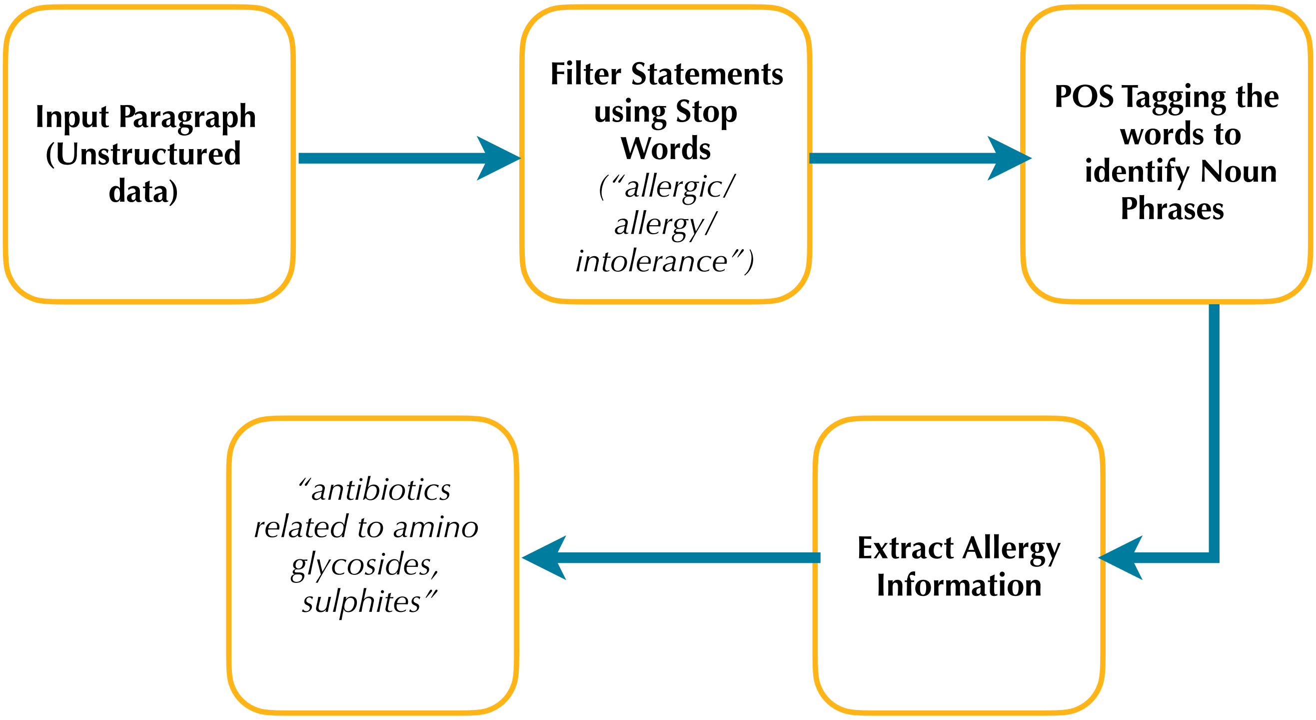
Knowledge Base Creation for Each Business Rule:

1. Drug Allergies:

Input unstructured data:

Do not take AMICIN 500MG INJECTION if you are allergic to AMICIN 500MG INJECTION, antibiotics related to aminoglycosides, sulphites, suffering from a disorder called myasthenia gravis (severe weakness of muscles in the body), and dehydration. Talk to your doctor before taking AMICIN 500MG INJECTION if you have any kidney problems, hearing problems such as tinnitus (buzzing or ringing in the ears). You are advised not to drive or operate heavy machinery after taking AMICIN 500MG INJECTION as it may cause dizziness. If you are pregnant or breast-feeding, talk to your doctor before taking AMICIN 500MG INJECTION. The doctor will prescribe you AMICIN 500MG INJECTION only if the benefits outweigh the potential risks.

Information Extraction Process:



Knowledge Base Information:

Generic name	Allergy_flag
ACECLOFENAC	ibuprofen, diclofenac, naproxen, ketorolac
AMIKACIN	antibiotics related to aminoglycosides, sulphites
AMIODARONE	iodine, amiodarone hydrochloride
ASPIRIN+CLOPIDOGREL	clopidogrel, aspirin, salicylates, anti-inflammatories (pain killers)
ASPIRIN+ROSUVASTATIN	aspirin and rosuvastatin

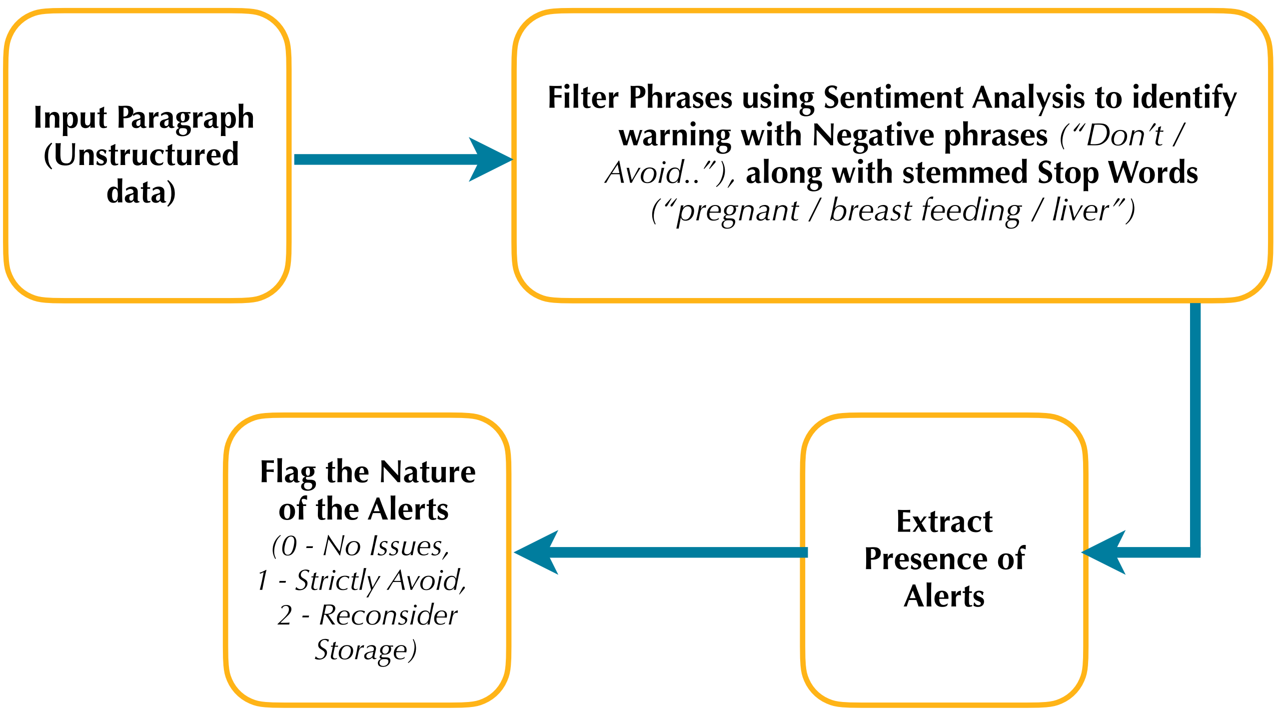
On pattern analysis, it is identified that the medicines sharing the same active ingredient, irrespective of medicine type *(tablets/ syrup)*, strength *(dosage can be varied at the time of administration)* have same allergy warnings. This is used during run-time to raise alerts.

2.Drug Contraindication:

Input unstructured data:

> Zerodol-Spas Tablet 10's can increase your risk of fatal heart attack or stroke. So, if you have had any recent heart surgery, do not use Zerodol-Spas Tablet 10's. Do not take Zerodol-Spas Tablet 10's if you are allergic to Zerodol-Spas Tablet 10's, have severe heart failure, have suffered bleeding problems such as bleeding from stomach or bowels while taking any pain killers or you have liver or kidney problems. Inform your doctor before taking Zerodol-Spas Tablet 10's if you have or had high blood pressure, heart problems, high cholesterol, diabetes, asthma, liver and kidney problems. Do not take Zerodol-Spas Tablet 10's if you are pregnant or breastfeeding unless prescribed. Zerodol-Spas Tablet 10's causes drowsiness and dizziness, so drive only if you are alert. Zerodol-Spas Tablet 10's should not be given to children as the safety have not been established. Avoid consuming alcohol along with Zerodol-Spas Tablet 10's as it could lead to increased drowsiness and can increase the risk of stomach bleeding. Stop taking Zerodol-Spas Tablet 10's and consult your doctor immediately if you have stomach pain or any signs of bleeding in intestine or stomach such as blood in stools. Do not take any other NSAID's (pain killer) for pain relief along with Zerodol-Spas Tablet 10's unless prescribed. Keep your doctor informed about your health condition and medicines to rule out any side-effects.

Information Extraction Process:



GenericName	Pregnancy_flag	Kidney_flag	Liver_flag	Breastfeeding_flag
THYROXINE SODIUM	2	2	2	2
ACEBROPHYLLINE+MENTHOL+TERBUTALIN	0	0	0	0
PROMETHAZINE	2	2	2	1
CHLORZOXAZONE+PARACETAMOL	2	2	1	2
TELMISARTAN	2	2	1	2

Knowledge Base Information:

On pattern analysis, it is identified that medicines sharing the same active ingredient irrespective of medicine type, strength have same caution flag across patient categories. This is used during run-time to raise alerts.

3.Therapeutic duplication/ Overdose:

Input unstructured data:

MEDICINENAME	Key ingredient	SUBSTITUTES
ACAMPROL TABLET	ACAMPROSATE	['ACAMPTAS 333MG TABLET']
AB PHYLLINE SR TABLET	ACEBROPHYLLINE	['ASCOVENT SR TABLET', 'AB FLO SR TABLET', 'BROPHYLE SR TABLET', 'ACEBROBID SR 200MG CAPSULE', "Cebocontin 200mg Cr Tablet 10's"]
Broclear Tablet 10's	ACEBROPHYLLINE+ ACETYLCYSTEINE	['ABIWAYS TABLET']

Knowledge base rule:

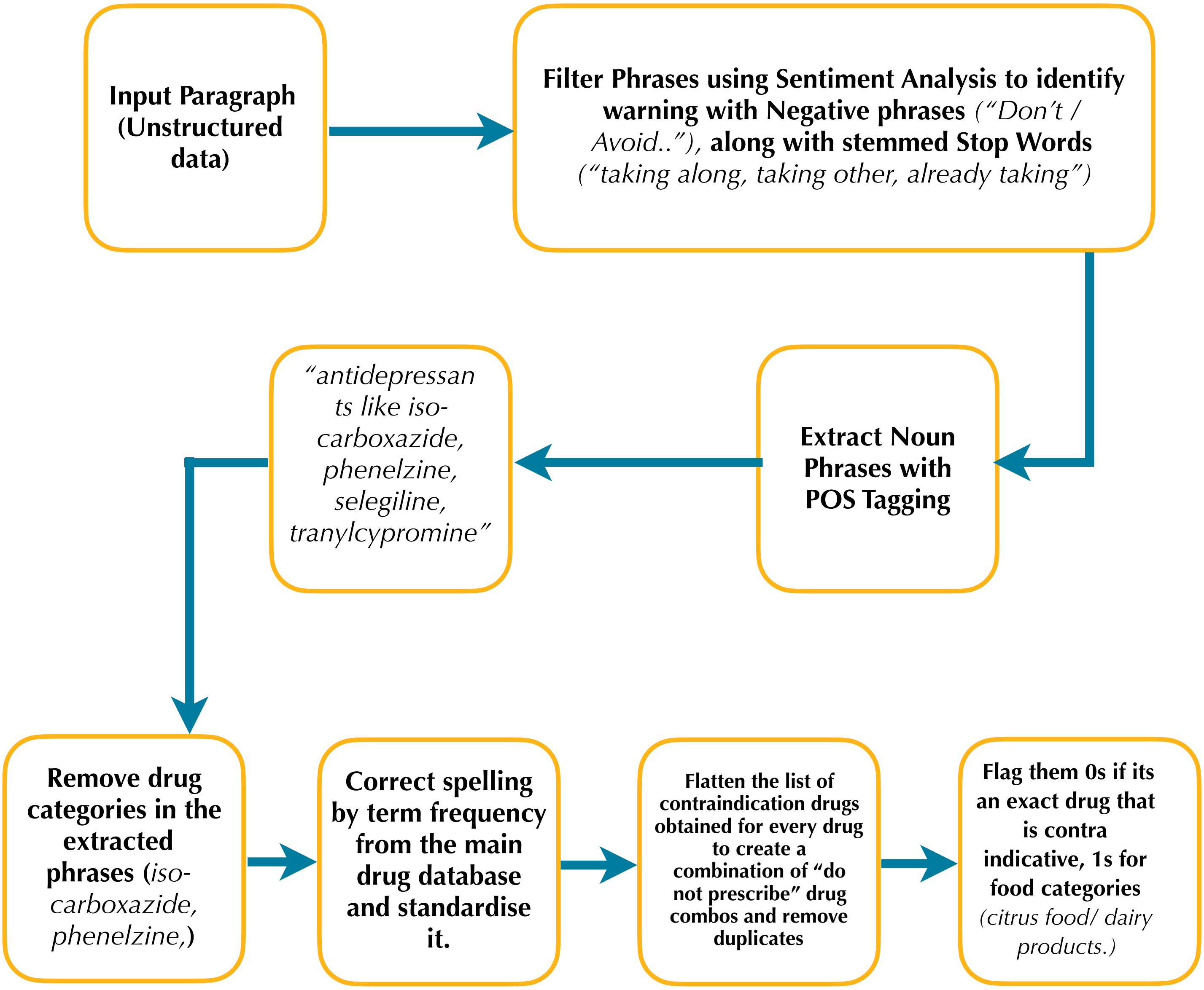
On pattern analysis, it is identified that medicines sharing the same active ingredient and medicine type (tablets/ syrup) irrespective of strength (dosage can be varied at the time of administration) can be substitutes to one another. Prescribing more than one substitutes on the same prescription or on consecutive prescription can lead to drug overdose. Another situation is that, when two drugs sharing a common active ingredient among other list of ingredients can lead to a therapeutic duplication which could be harmful at times. This active ingredient repetition is identified on run-time and the alerts are raised.

4.Drug Interaction:

Input unstructured data:

Directions
Do not take Tryptomer 10 mg Tablet 30's if you are allergic to any of its contents if you recently had a heart attack, or heart problems, are taking other antidepressants like isocarboxazid, phenelzine, selegiline, tranylcypromine or have taken them in the last 14 days, have severe liver disease or if you have taken moclobemide (used to treat depression and social anxiety). Inform your doctor before taking Tryptomer 10 mg Tablet 30's if you have heart rhythm problems or hypotension. Consult your doctor immediately if you have suicidal thoughts such as killing or harming yourself. Do not take Tryptomer 10 mg Tablet 30's if you are pregnant or breastfeeding unless prescribed. Tryptomer 10 mg Tablet 30's causes drowsiness and dizziness, do not drive unless you are alert. Tryptomer 10 mg Tablet 30's can be given to children above 6 years if prescribed by the doctor for treating bedwetting. Tryptomer 10 mg Tablet 30's should not be given to children for treating depression or neuropathic pain. Avoid consuming alcohol along with Tryptomer 10 mg Tablet 30's as it could lead to increased drowsiness and dizziness. Rise slowly from sitting or lying position as Tryptomer 10 mg Tablet 30's causes orthostatic hypotension (sudden lowering in blood pressure leading to dizziness on standing).

Information Extraction Process:



Drug1	Drug2	DI_Code
ACECLOFENAC+DROTAVERINE	IBUPROFEN	0
ACECLOFENAC+DROTAVERINE	DICLOFENAC	0
ACECLOFENAC+DROTAVERINE	NAPROXEN	0
ACECLOFENAC+DROTAVERINE	KETOROLAC	0
ACECLOFENAC+PARACETAMOL+THIOCOLCHICOSIDE	IBUPROFEN	0
ACECLOFENAC+PARACETAMOL+THIOCOLCHICOSIDE	DICLOFENAC	0

Knowledge Base Information:

On pattern analysis, it is identified that certain drugs cause harm when taken along a certain ingredient compound. Eg. Nitrates. On such scenarios, alert must be raised for all drugs containing nitrates Eg. Glyceryl DiNitrate) when prescribed along any of the red alert drug list.

Building Apriori Thresholds from the Database

Part I: Pre-Processing

i. The list of unique generic names are obtained from the indenting data sheet.

ii. The generic names are cleaned to extract the active ingredient in the drug, by removing the strength information and standardising them using Regular Expressions.
For example: ASPIRIN 40 MG , ASPIRINE 40 MG, ASPIRIN, ASPIRIN 30 MG => ASPIRIN

iii. The corrected generic names are mapped back to the indents.

IPNUMBER	MEDICINENAME	DRUGCODE	GENERICNAME	Corrected_GenericNames
[redacted]	THYROX 50MCG TAB	THY0005	THYROXINE SODIUM	THYROXINE SODIUM
[redacted]	PULMOCLEAR SYP 100ML	PUL0108	ACEBROFYLINE+GUAIPHENESIN+TERBUTALINE	ACEBROPHYLLINE+MENTHOL+TERBUTALINE
[redacted]	PHENERGAN 100ML SYP	PHE0027	PROMETHAZINE	PROMETHAZINE
[redacted]	MYOSPAZ TAB	MYO0002	PARACETAMOL 325MG, CHLORZOXAZONE 250MG	CHLORZOXAZONE+PARACETAMOL
[redacted]	TELLZY 40MG TAB	TEL0062	TELMISARTAN	TELMISARTAN

Part II: Fetching Item Sets based on Frequency

i. The top 40% of generic names prescribed the most are filtered out.

ii. All indents containing the top 40% generic names are then filtered out.

iii. The duplicate records are removed, with IP number as the filter.

	IPNUMBER	Corrected_GenericNames
0	[redacted]	THYROXINE SODIUM
5	[redacted]	LEVETIRACETAM
8	[redacted]	METOPROLOL
17	[redacted]	IPRATROPIUM BROMIDE+LEVOSALBUTAMOL
18	[redacted]	METOPROLOL
...	[redacted]	...
1313182	[redacted]	PARACETAMOL
1313183	[redacted]	CEFTRIAXONE
1313185	[redacted]	ONDANSETRON
1313186	[redacted]	PANTOPRAZOLE
1313189	[redacted]	PARACETAMOL

iv. The list of generic names are then grouped by IP and flattened, and converted into a pivot table.

Corrected_GenericNames	ACETYLCYSTEINE	ASCORBIC ACID+BIOTIN+CYANOCOBALAMIN+FOLIC ACID+NICOTINAMIDE+PYRIDOXINE+RIBOFLAVIN+THIAMINE+VITAMIN B5	ASPIRIN	ATORVASTATIN	BENZYL ALCOHOL+DICLOFENAC SODIUM	BUDESONIDE
0	False	False	False	False	True	False
2	False	False	False	False	False	False
3	False	False	False	False	False	False
8	True	True	False	False	False	True
10	False	False	False	True	False	True
...	...	...	...	...	...	...
27203	False	False	True	True	False	False
27204	False	False	False	True	False	False
27205	False	False	False	False	False	False
27206	True	False	False	True	False	False
27207	False	False	False	False	False	False

v. The Apriori algorithm is run with a support threshold of 0.1 via Google Colab.

support	itemsets
0.6139474	PANTOPRAZOLE, PARACETAMOL
0.436114086	PANTOPRAZOLE, ONDANSETRON
0.350668412	ONDANSETRON, PARACETAMOL
0.321750169	PANTOPRAZOLE, ONDANSETRON, PARACETAMOL
0.256813043	PANTOPRAZOLE, CEFOPERAZONE+SULBACTAM
0.225276687	CEFOPERAZONE+SULBACTAM, PARACETAMOL
0.212662144	PANTOPRAZOLE, LIDOCAINE
0.208695307	PANTOPRAZOLE, BENZYL ALCOHOL+DICLOFENAC SODIUM
0.20472847	PANTOPRAZOLE, ENOXAPARIN
0.203419414	PANTOPRAZOLE, CEFOPERAZONE+SULBACTAM, PARACETAMOL
0.184576937	LIDOCAINE, PARACETAMOL
0.17743663	BENZYL ALCOHOL+DICLOFENAC SODIUM, PARACETAMOL

vi. The wards contributing to the top 50% indents are identified with the help of EDA

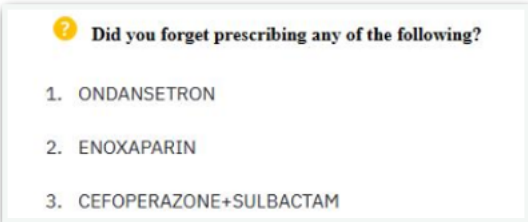
Group 3: Ward Names with > 1 Lakh Indents (53% of all indents)

Ward Name	Total Indents
int 2floor	114,430
Int Block 3rd Floor	121,108
1st FloorSpecial Ward	123,022
3rd FloorSpecial Ward	139,444
2nd FloorSpecial Ward	196,291

vii. The indents are now filtered with these set of wards to create individual subsets.
viii. The pivot table is regenerated within each ward and Apriori algorithm is run again.
ix. Ward specific frequent item sets of drugs are identified.

Part III: Knowledge Base Creation & Auto - Tuning

i. Unique prescription patterns that are not distinctly identified in the overall Apriori results are identified.
ii. While suggesting missing drugs, ward specific Apriori results are prioritised, if the ward mentioned is among the top 40% indent contributors.
iii. The support value of ward specific frequent item sets are incremented with a higher weightage to top the suggestion list at runtime.



iv. This top 40% result tuning would be automated in upcoming versions of application along with Apriori algorithm being re-run as a timed batch process to update the knowledge base.

Benefits & Caveats of the Decision Support System

Benefits of the DSS

i. The decision support system built is **scientific and uses data as its backbone for operation.** This ensures that all alerts raised are not opinion based, and are completely defined by predefined medical rules and historical indenting patterns in the hospital.

ii. Since most of the medication errors that occur at Apollo Hospitals are due to prescription errors, it is prudent to exercise precaution at the very first stage. **This system is built to alert medical professionals as early as possible in the medication process, mainly because time is of utmost essence in preventing an error from reaching the patient.**

iii. The entire process being 100% electronic implies that no error or potential future error goes unnoticed. This, combined with our decision support system's account of overridden alerts would help Apollo hospitals **exercise more control over the medication process.**

iv. The quality team of Apollo, who would serve as superusers of the system would get insights about the **performance of each individual user indenting over a period of time at the hospital unit.** This could in turn help the **quality team to assess high risk areas, suggest targeted improvement processes to prevent excess overriding of alerts**, indirectly fine-tuning the medication process.

v. Since the system predominantly works at the level of drug monitoring at the time of prescription / indenting, the quality team can also **constantly monitor those drugs** which potentially raise more alerts. They could subsequently workout remedial measures to **prevent errors arising due to the use of these drugs by looking at alternatives which are less prone to cost error, in consultation with doctors.**

vi. If an error were to occur despite the system suggesting alerts, **the quality team at Apollo would be able to backtrack the error to the exact user who overrode the alert, and the indent that was overridden. This would increase accountability in the medication process, owing to data traceability.**

vii. This decision support system, if used as intended, would help streamline and tighten the medication process and help save the time of those involved in indenting, thereby **increasing productivity.**

viii. **The DSS, along with the tentative data collection mechanisms recommended, when implemented, would serve as the first step in the digital transformation of the medication process at Apollo hospitals, paving way for future implementation of more advanced machine learning and artificial intelligence-based technology to help redefine the quality of health care provided at Apollo hospitals.**

ix. **Finally, more than anything else, reducing the probability of occurrence of even one medication error can potentially save a human life, and that is the biggest impact this system would have on the society as a whole.**

Caveats of the DSS

i. The DSS only considers drugs being prescribed as the prima facie contributor to errors, but as explored earlier, the FishBone analysis showed multiple factors that could be featured. More features in the future would imply more error alerting effectiveness.

ii. The list of drugs used in the construction of this system is not 100% exhaustive, given that not all drugs possible could have been indented in 2019, nor does the Apollo Pharmacy website and the formulary account for every single drug.

iii. The system is backed by the indenting data insights of one specific hospital, and this is completely unique to each hospital. As a result, this DSS can not be instantly scaled up for other hospital units. However, proper data collection, cleaning and preprocessing of indents in the other units can assist in the replication of this system across the board.

iv. The system assumes the trustworthiness and authenticity of the data provided in the Apollo Pharmacy website, and could potentially throw alerts which are correct from a data standpoint, but might make lesser medical sense.

v. Apollo Hospitals would have to periodically update the system with fresher set of indents, which would reduce the redundancy of the issue of newer drugs in the future, and increase the efficacy of the alerts being recommended.

Conclusion & Recommendations

Proposed Solutions

Decision Support System

The decision support system designed and explained in the previous section(s) would serve as an apt tool to accomplish the following objectives:

- Provide Apollo Hospitals' quality team to better control the medication process.
- Give a holistic account of medication error tracing - a bird's eye view of all the errors in a time period + the detailed performance of each indenting individual separately.
- Get Apollo Hospitals to monitor the impact of individual / combination of drugs on error occurrences.
- Data traceability provides an option to backtrack future errors and pinpoint the indenting activity that caused it.
- When deployed, act as the catalyst to drive digital transformation of the medication process at Apollo Hospitals, and
- Reduce the occurrence of overall errors at Apollo Hospitals, potentially saving lives.

Recommended Features for Future Data Collection

Upon deep diving into the medication process using FishBone analysis, several factors were identified that impacted medication errors. Capturing these extra data points and features over a period of time would be vital for building a more advanced AI-ML based system in the future for Apollo Hospitals. A few of the features are listed below

Variable / Feature	Description	Rationale
Patient_Age	Age of the patient	For drug dosage calculations
Patient_Height	Height of the patient	For drug dosage calculations
Patient_Weight	Weight of the patient	For drug dosage calculations
Patient_Origin	State / city of origin	To check for language barriers
Patient_Specialty	Chief complaint specialty	For error area classification
Med_Name_Old	Original Prescribed Medicine	For error tracing in substitutions
Dosage_Old	Original Prescribed Dosage	For error tracing in substitutions
Med_Name_New	Changed Medicine Name	New replaced drug's data
Dosage_New	Changed Medicine Dosage	New replaced drug's data
Ward_Bed_Count	Size of the ward	For caregiver vs patient ratios
Ward_Bed_Position	Position of bed in the ward	Relative position to nurse station
Ward_Percent_Occupancy	% Occupancy	For caregiver vs patient ratios
ICU_Bed_Count	Size of the ICU	For caregiver vs patient ratios
ICU_Bed_Position	Position of bed in the ICU	Relative position to nurse station

Variable / Feature	Description	Rationale
ICU_Percent_Occupancy	% Occupancy	For caregiver vs patient ratios
Nurse_Join_Date	Date of Joining of Apollo	For checking familiarity with SOPs
Nurse_Prior_Exp	Prior Experience	Error Process Tracing
Nurse_Shift_Time_In	When nurse begins duty	Shift timings & duration vs Error
Nurse_Shift_Time_Out	When nurse ends duty	Shift timings & duration vs Error
Nurse_Shift_Type	Normal vs Extended Shifts	Shift timings & duration vs Error
Nurse_Ward_Info	Nurse's ward responsibility	For caregiver vs patient ratios
Nurse_Age	Age of the Nurse	Error Process Tracing
Doc_Join_Date	Date of Joining of Apollo	For checking familiarity with SOPs
Doc_Prior_Exp	Prior Experience	Error Process Tracing
Doc_Shift_Time_In	When doctor begins duty	Shift timings & duration vs Error
Doc_Shift_Time_Out	When doctor ends duty	Shift timings & duration vs Error
Doc_Shift_Type	Normal vs Extended Shifts	Shift timings & duration vs Error
Doc_Age	Age of the Doctor	Error Process Tracing
Doc_Type	Resident vs Consultant	Prescription comparison tracing
Doc_Patient_Count	Number of consultations /round patients visited	For caregiver vs patient ratios
PC_Age	Age of the Pharmacist	Error Process Tracing
PC_Exp	Prior Experience	Error Process Tracing
PC_Count	No. of pharmacists on duty	Stress levels during shift
PC_Join_Date	Date of Joining of Apollo	For checking familiarity with SOPs
PC_Shift_Time_In	When pharmacist begins duty	Shift timings & duration vs Error
PC_Shift_Time_Out	When pharmacist ends duty	Shift timings & duration vs Error
Drug_Dis_Time	Time of Dispensing Drug	Stress levels during shift
Drug_Hr_Count	Hourly amounts of drugs dispensed	Stress levels during shift

Note: It is also advised to digitise the drug order sheet; use DD/MMM/YYYY date type; use standardised abbreviations, error types, and error causes; have a single database instead of multiple sources of truth; and characterise drug categories by different therapeutic levels (0,1,2..) for minimising common error data collection mistakes.

Recommendations

i. This DSS would assist doctors, nurses and other medical professionals to make more informed decisions during the medication process. In no ways is this system intended to replace the instincts of these personnel, and therefore it is to be noted that the expertise and discretion of these professionals would remain supreme.

ii. The system is currently independent of the patients personal medical records and history. Integration with it would vastly enhance the capabilities of the system to be intuitive and robust in suggesting alerts for medications.

iii. As mentioned before, drugs in the prescription alone drive a majority of the analytics behind this system. However, since there are far many factors that can contribute to the occurrence of a medication error, it would be essential for Apollo Hospitals to collect data across these features in order to get a more holistic reasoning behind medication errors. More features mean more insights; more insights mean more power to the DSS, and more power to the DSS means better prevention of medication errors.

iv. As organisations across the world transform digitally, it is essential for Apollo Hospitals too, to create, maintain and nourish their data pipelines constantly, not just for medications, but across all processes that are linked with the service quality. This would open a lot of doors for more AI & ML based analytics to be implemented in the future, all the while providing Apollo Hospitals with enhanced capabilities to continue their mission - achieving and maintaining excellence in healthcare for the benefit of humanity.

Bibliography

Reference Papers, Links & Resources:

1. Campino A., Lopez-Herrera M. C., Lopez-de-Heredia I., Valls-i-Soler, A. Medication errors in a neonatal intensive care unit: Influence of observation on the error rate, Acta Paediatrica, 2008; 97:1589–1594.
2. Woolf S. H., Kuzel A. J., Dovey S. M., Phillips R. L. A String of Mistakes: The Importance of Cascade Analysis in Describing, Counting, and Preventing Medical Errors. Annals of Family Medicine. 2004; 2(4): 317-326.
3. Anderson J. G., Jay S. J., Anderson M., Hunt T.J. Evaluating the Potential Effectiveness of Using Computerized Information Systems to Prevent Adverse Drug Events. AMIA Annual Symposium 1997: 228-232.
4. Halpern M.T., Brown R.E., Revicki D. A., Togias A.G. An example of using computer simulation to predict pharmaceutical costs and outcomes. Winter Simulation Conference Proceedings 1994: 850-855.
5. Koppel R. et al., "Workarounds to Barcode Medication Administration Systems: Their Occurrences, Causes, and Threats to Patient Safety," Journal of the American Medical Informatics Association 2008; 15 (4):408–423
6. Bates D. W, Cohen M, Leape L. L, et al. Reducing the frequency of errors in medicine using information technology. J Am Med Inform Assoc. 2001; 8: 299–308.
7. Leape, L. L., Bates, D. W., Cullen, D. J., Cooper, J., Demonaco, H. J., Gallivan, T., et. al. Systems analysis of adverse drug events. ADE Prevention Study Group. JAMA. 1995; 274(1):35–43.
8. Phillips, J., Beam, S., Brinker, A., Holquist, C., Honig, P., Lee, L., Parner, C. Retrospective analysis of mortalities associated with medication errors. Am. J. Health Sys. Pharm. 2001; 58(19):1835–1841.
9. Lesar, T. S., Briceland, L., and Stein, D. S., Factors related to errors in medication prescribing.
10. Jia P, Zhang L, Chen J, Zhao P, Zhang M. The Effects of Clinical Decision Support Systems on Medication Safety: An Overview. *PLoS One*. 2016;11(12):e0167683. Published 2016 Dec 15. doi:10.1371/journal.pone.0167683
11. https://www.fda.gov/regulatory-information/fdaaa-implementation-chart/usp-therapeutic-categories-model-guidelines
12. https://www.who.int/health-topics/hospitals
13. https://www.fda.gov/drugs/information-consumers-and-patients-drugs/working-reduce-medication-errors
14. https://streamlit.io/
15. https://www.researchgate.net/publication /343194021_Role_of_Artificial_Intelligence_in_Patient_Safety_Outcomes_Systematic_Literature_Review
16. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2464935/
17. https://news.microsoft.com/life/videos/prevent-medication-errors/
18. https://journals.lww.com/md-journal/fulltext/2020/05010 association_rule_mining_for_the_ordered_placement.57.aspx
19. https://www.apollohospitals.com
20. https://www.apollopharmacy.in/
21. https://www.webmd.com
22. https://rxlist.com
23. https://drugs.com
24. Asdada https://medscape.com
25. https://www.ibm.com/docs/es/db2/9.7?topic=SSEPGG_9.7.0/com.ibm.datatools.datamining.doc c_ta_information_extraction.html

Tools used: MS Excel, Tableau, Jupyter Notebook, WebDriver v86, Selenium & Beautiful Soup Packages for Python, R-Studio, IBM SPSS Statistics, Rattle, Adobe Photoshop, Pages