Hospital Readmission Risk Prediction

Hospital readmission risk is the probability that a patient will return to the hospital within a predetermined time frame following their original discharge. Readmissions to the hospital can be expensive for both patients and the healthcare system, and they may be a sign of underlying problems with the standard of treatment or patient management.

Hospital readmission risk can be affected by a number of variables, such as the patient's medical history, the seriousness of their sickness or condition, and any comorbidities they may have. Patient demographics, such as age, gender, and race, as well as social and environmental factors, such as housing and access to healthcare, may also influence the likelihood of readmission.

Problem Statement

An AI-powered platform that analyses a patient's medical history and present state to forecast the possibility that they will need to be readmitted to the hospital after being discharged.

Market/Customer/Business Need Assessment

Market Need: Hospital readmission risk prediction technologies that might assist healthcare professionals in lowering patient readmission rates are increasingly needed in the market. These solutions are in demand because of things like rising healthcare expenditures, a focus on patient-centered treatment, and the requirement for better resource management.

Customer Need: Customer Need: To avoid readmission, patients who are released from hospitals frequently need continuous care and assistance. Healthcare professionals can deliver focused interventions to reduce readmission risk and guarantee that patients receive the care they require by properly assessing a patient's risk of readmission. This can increase patient outcomes and satisfaction while lowering overall healthcare costs.

Business Need: Hospital readmissions can be expensive for healthcare providers, hence many hospitals and healthcare organisations prioritise lowering readmission rates. Healthcare providers may lower readmission rates, enhance patient outcomes, and optimise resource allocation by using an AI-powered hospital readmission risk prediction solution. This can result in reduced costs, better care quality, and a competitive edge in the healthcare sector.

Target Specifications and Characterization

Healthcare Providers: Hospitals, clinics, and other healthcare organisations are the main target market for Hospital Readmission Risk prediction tools. These clients are seeking solutions that will enable them to lower readmission rates, enhance patient outcomes, and efficiently use resources.

Care Managers: Potential clients for Hospital Readmission Risk prediction tools include care managers who are in charge of patient care coordination, discharge planning, and follow-up care. These clients are searching for products that would enable them to recognise high-risk patients and create individualised care plans to lower the chance of readmission.

Patients and Caregivers: Despite not being direct customers of Hospital Readmission Risk prediction technologies, patients and carers can nonetheless gain from their deployment. Healthcare professionals can better support and care for patients and enhance their general health outcomes by lowering the likelihood of readmission.

Healthcare Payers: Potential clients for Hospital Readmission Risk prediction technologies include healthcare payers like insurance firms and governmental organisations. These clients are seeking solutions that would enable them to save healthcare expenses while raising the standard of care for their patients or beneficiaries.

External Search (online information sources/references/links)

I used Predicting Hospital Readmissions dataset from Kaggle. The dataset consists of 17 attributes and 25000 data points.
The dataset looks like below:

```
df.head(10)
```

	age	time_in_hospital	n_lab_procedures	n_procedures	n_medications	n_outpatient	n_inpatient	n_emergency	medical_specialty	diag_1	diag_2
0	[70-80)	8	72	1	18	2	0	0	Missing	Circulatory	Respiratory
1	[70-80)	3	34	2	13	0	0	0	Other	Other	Other
2	[50-60)	5	45	0	18	0	0	0	Missing	Circulatory	Circulatory
3	[70-80)	2	36	0	12	1	0	0	Missing	Circulatory	Other
4	[60-70)	1	42	0	7	0	0	0	InternalMedicine	Other	Circulatory
5	[40-50)	2	51	0	10	0	0	0	Missing	Other	Other
6	[50-60)	4	44	2	21	0	0	0	Missing	Injury	Other
7	[60-70)	1	19	6	16	0	0	1	Other	Circulatory	Other
8	[80-90)	4	67	3	13	0	0	0	InternalMedicine	Digestive	Other
9	[70-80)	8	37	1	18	0	0	0	Family/GeneralPractice	Respiratory	Respiratory

Research Papers and Journals: I went through various research papers to get a good understanding of the prediction using ML. One of the research papers was Effective hospital readmission prediction models using machine-learned Features. (doi - https://doi.org/10.1186/s12913-022-08748-y)

Benchmarking

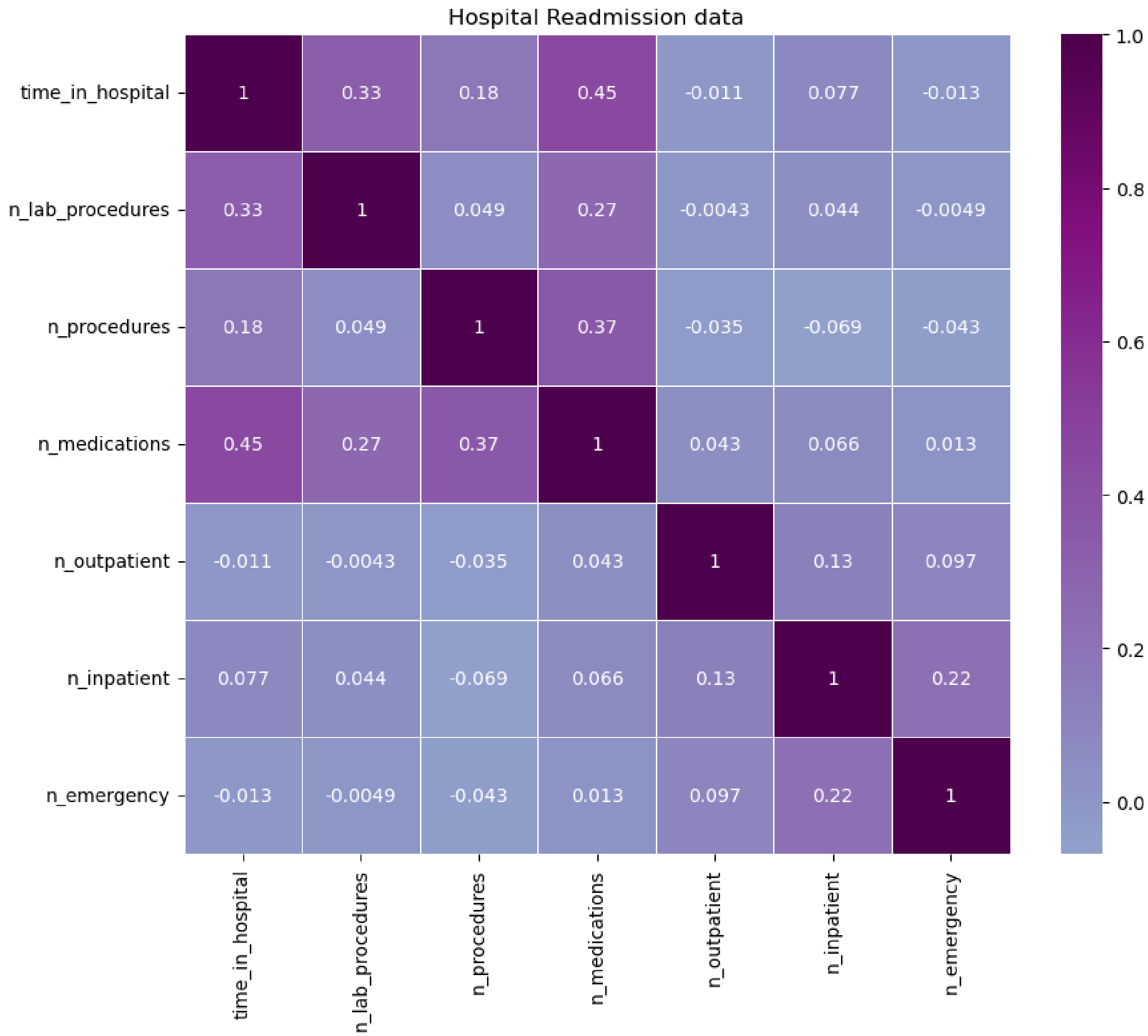
A crucial element in creating a successful Hospital Readmission Risk prediction method is benchmarking competing goods and services. You can find market gaps and potential areas for improvement by contrasting and comparing current goods and services.

When benchmarking alternate products or comparing with existing products/services for a Hospital Readmission Risk prediction, it is essential to consider the following factors:

1. Accuracy: The accuracy of the predictive model is critical in assessing the risk of hospital readmission. It is important to benchmark the accuracy of the predictive model against existing solutions in the market.
2. Scalability: The predictive model should be able to scale with increasing volumes of data and patients. It is important to benchmark the scalability of the solution against existing solutions in the market.
3. Cost: Cost is a crucial factor in healthcare solutions. It is important to benchmark the cost of the predictive model against existing solutions in the market.

4. Integration: The predictive model should be easily integrated with the existing hospital information systems. It is important to benchmark the integration capabilities of the solution against existing solutions in the market.
5. Interpretability: The predictive model should be easy to interpret for healthcare providers and patients. It is important to benchmark the interpretability of the solution against existing solutions in the market.

Correlation Matrix For Hospital Readmission



Applicable Patents

It is important to conduct a thorough patent search to ensure that the Hospital Readmission Risk prediction project does not infringe on any existing patents.

Some of the relevant patents in this field include:

1. U.S. Patent No. 10,620,414: System and method for predicting hospital readmissions using machine learning
2. U.S. Patent No. 9,652,684: Method and system for predicting hospital readmission risk using electronic medical records
3. U.S. Patent No. 9,452,938: Predictive model for hospital readmission risk based on medical history and social determinants of health
4. U.S. Patent No. 9,233,090: Method for predicting readmissions for patients with heart failure using machine learning techniques
5. U.S. Patent No. 9,187,574: Predictive modelling of hospital readmissions using electronic medical records

Applicable Regulations

In India, there are several regulations that apply to the healthcare industry, including regulations related to patient privacy and data protection. The following are some of the applicable regulations that may be relevant to Hospital Readmission Risk Prediction:

1. Personal Data Protection Bill, 2019: This bill, which is currently being debated in the Indian Parliament, aims to regulate the collection, use, and storage of personal data. The bill imposes strict obligations on organizations that collect personal data and includes provisions related to data security and breach notification.

2. Health Insurance Portability and Accountability Act (HIPAA): Although HIPAA is a US law, it may apply to Indian companies that have business dealings with US healthcare organizations. HIPAA imposes strict requirements on the handling of protected health information (PHI) and requires organizations to implement appropriate safeguards to protect the privacy and security of PHI.

3. Indian Medical Council (Professional Conduct, Etiquette and Ethics) Regulations, 2002: These regulations govern the professional conduct of

doctors and other medical professionals in India. They include provisions related to patient privacy and confidentiality.

4. Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules, 2011: These rules set out the requirements for the collection, use, and storage of sensitive personal information, including medical records.

Applicable Constraints

The Hospital Readmission Risk Prediction project may have several constraints that need to be considered. Some of the common constraints are:

Physical space may be needed for the project to put up the necessary hardware and software infrastructure. Servers, storage units, and other machinery may be included.

Budget: The project might necessitate a substantial expenditure in terms of manpower, technology, and software. Depending on the project's complexity and the level of skill required, the price may change.

Knowledge: The project may call for knowledge in a variety of fields, such as data analytics, machine learning, and healthcare industry-specific knowledge. Depending on the project's size, it can be required to work with or employ professionals in various fields.

Data accessibility: Data accessibility, data privacy, and data quality can all pose serious challenges. Getting the necessary information from patients or healthcare professionals might be difficult.

Business Model

The Hospital Readmission Risk Prediction project can have several potential business models, some of which are:

Model-based on subscriptions: In this model, hospitals or other healthcare organisations can subscribe to the platform. The subscription cost may be determined by the total number of patients or the anticipated readmissions.

Pay-per-use approach: In this approach, the number of anticipated readmissions will determine how much hospitals or other healthcare organisations would pay for the platform. A per-patient or per-episode charge structure is an option.

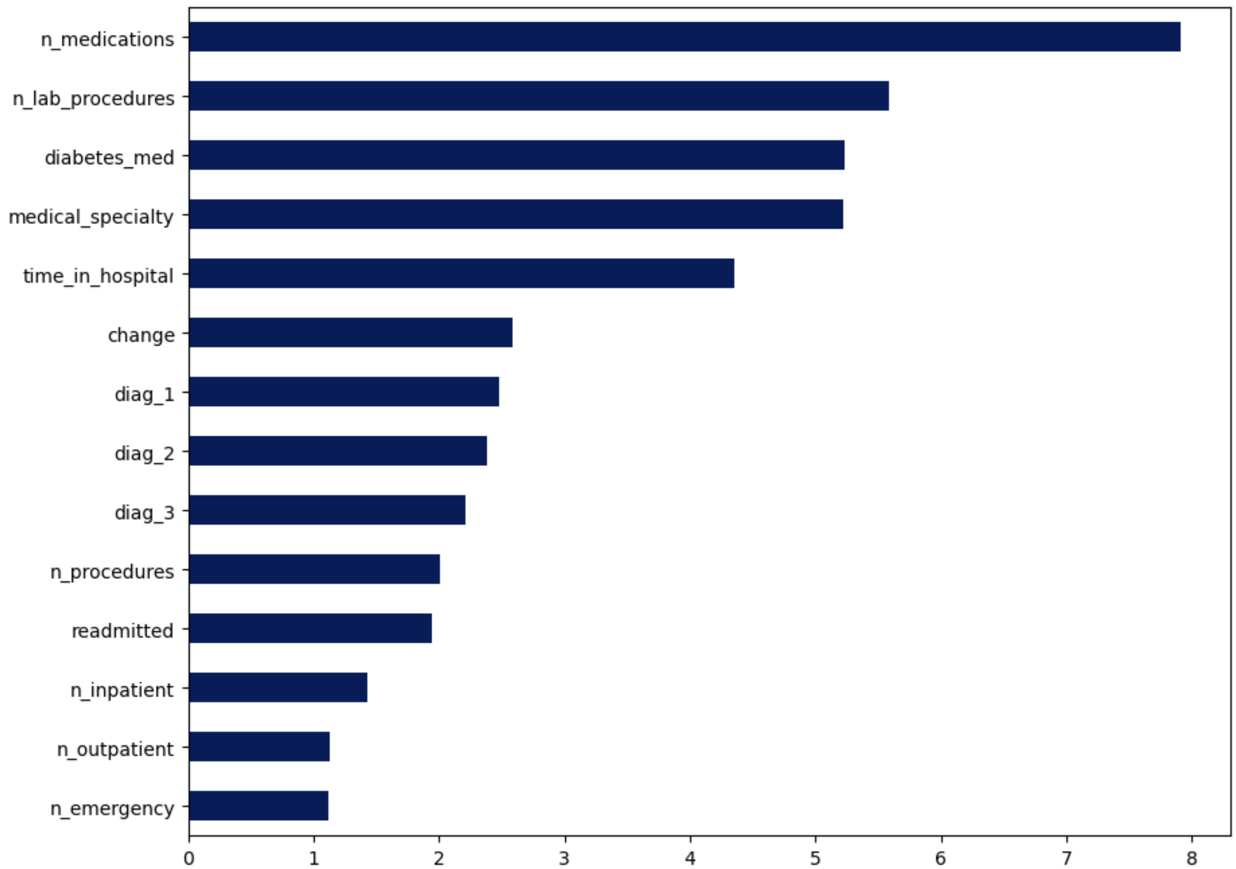
Approach for Sharing Revenue: In this approach, the platform may be provided without charge to hospitals or other healthcare organizations, with revenue shared among the parties. Based on the cost savings achieved by fewer readmissions, the revenue might be divided.

Model for Consulting: In this model, the project team can give hospitals or other healthcare organisations consulting services to assist in implementing the platform and integrating it with their current systems. The cost may vary depending on the intricacy of the implementation or the number of consulting hours.

Concept Generation

To meet our needs, this product necessitates the creation of machine learning models. It's less intimidating to modify existing models for our needs than to build them from start. A well-trained model can be constructed or repurposed. However, creating a model using the available tools and data is slow but doable. The analyst could choose to input data as quickly as possible. It will take some work to get this precision because it is dangerous to rely only on the Classic Machine Learning method.

First, we clean the data and find insights from the data through various visualisations. Selecting the necessary features is an important factor which can be done through the Variance Inflation Factor(VIF). The variance inflation factor calculates how much an independent variable's behaviour (variance) is inflated by its interaction and association with other independent variables. Also, we can check the multi-collinearity among the variables to choose the variables.
The below bar chart shows how each variable is multi-co-linear with each variable. Therefore, each of the features is independent of each other.



To train the data, we split the data into training and testing with a ratio of 8:2 respectively.

```
In [119]: x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.2)
```

```
In [120]: # Train the logistic regression model
          model = LogisticRegression()
          model.fit(x_train, y_train)
```

We have done logistic regression, random forest and gradient boosting to find the best model.
The gradient boosting model gave us the best results. The confusion matrix was plotted for the same.

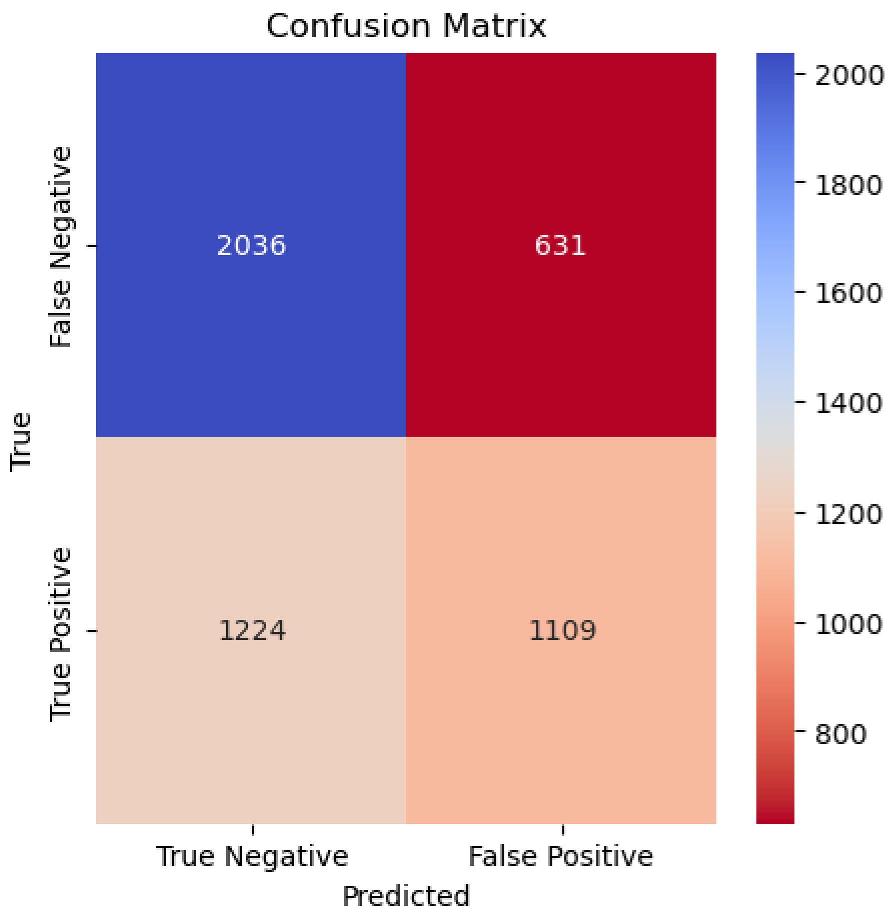
```
pred=gbc.predict(xgb_test)
print(classification_report(ygb_test, pred))
```

```
              precision    recall  f1-score   support

           0       0.62      0.76      0.69      2667
           1       0.64      0.48      0.54      2333

    accuracy                           0.63      5000
   macro avg       0.63      0.62      0.62      5000
weighted avg       0.63      0.63      0.62      5000
```

The evaluation metrics for the gradient boosting model and the confusion matrix.



To further improve our model, we used hyperparameter tuning to make the model more accurate.

Random search technique for hyperparameter tuning

```
In [159]: from scipy.stats import uniform
          from sklearn.model_selection import RandomizedSearchCV
```

```
In [160]: # Define the hyperparameter search space
          param_dist = {'learning_rate': uniform(0.01, 1.0),
                        'n_estimators': range(50, 500),
                        'max_depth': range(1, 11),
                        'min_samples_split': range(2, 21),
                        'min_samples_leaf': range(1, 21),
                        'subsample': uniform(0.01, 1.0)}

          # Perform random search with 5-fold cross-validation
          random_search = RandomizedSearchCV(gbc, param_distributions=param_dist, n_iter=100, cv=5, n_jobs=-1)
          random_search.fit(xgb_train,ygb_train)

          # Print the best hyperparameters found
          print("Best parameters:", random_search.best_params_)
```

```
In [161]: accuracy = accuracy_score(ygb_test, pred)
```

```
In [162]: print('Accuracy:',accuracy)

          Accuracy: 0.629
```

Concept Development

Clearly defining the issue or opportunity that the project is attempting to address is the first stage. In this situation, the issue or opportunity is the high rate of hospital readmissions and the requirement for a tool that can forecast the chance of readmission.

Conduct market research to determine trends in the healthcare sector, competition requirements and preferences, and the customer wants and preferences.

Create a list of concepts and ideas that could be used for the project, such as hybrid models, mobile applications, real-time monitoring, collaborative platforms, and patient engagement and education.

Ideas for improvement should be evaluated based on their viability, cost, scalability, and effect on patient outcomes.
Pick the winning concept: Based on the review, choose the best concept and develop it further.

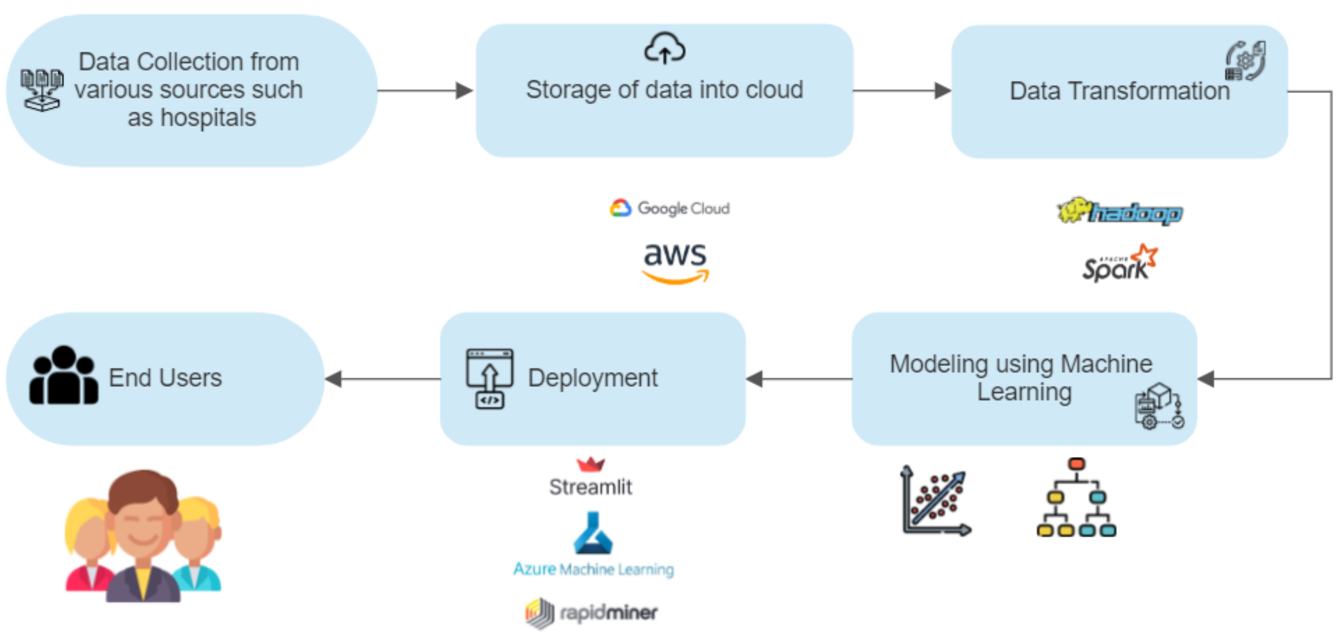
Create a prototype: Create a working model of the chosen concept to evaluate its viability and effectiveness.

Test the prototype: To gather feedback and find any areas that need improvement, test the prototype on a small number of users.

Refine the concept: Adjust the concept in light of user input and additional testing.
Complete the idea: Complete the idea and get ready to implement it.

Implement the idea: Put the idea into practice while regularly assessing its performance to find areas for development and making the required adjustments.

Final Product Prototype



There will be a front-end and back-end in our prototype. The back-end consists of data collection, data storage and maintenance, data transformation, modelling of data and backend coding for deployment. The front end contains the user interface through which various users such as patients, caretakers and doctors could access the conditions and medications required. Therefore an easy accessibility to doctors is also possible through this platform.

Code Implementation

Some Basic Visualizations on Real World or Augmented Data, Simple EDA and a ML Modelling are done and uploaded to the GitHub repository 🔗.

Conclusion

In conclusion, the creation of an AI-powered platform for predicting hospital readmissions is a useful tool that can assist healthcare professionals in identifying patients who are at a high risk of being readmitted and offering them the proper therapies to lower that risk. The platform analyses patient data using machine learning algorithms and provides real-time predictions, assisting healthcare professionals in making decisions that will enhance patient outcomes.

The potential benefits for patients, healthcare providers, and healthcare systems as a whole outweigh the hurdles of deploying such a platform, including regulatory compliance, data privacy issues, and the requirement for sufficient resources and experience.

The Hospital Readmission Risk Prediction platform has the ability to revolutionise the way healthcare is provided by offering individualised care that is suited to each patient's particular needs with more development and improvement.