



MACHINE LEARNING MSIB BISA AI

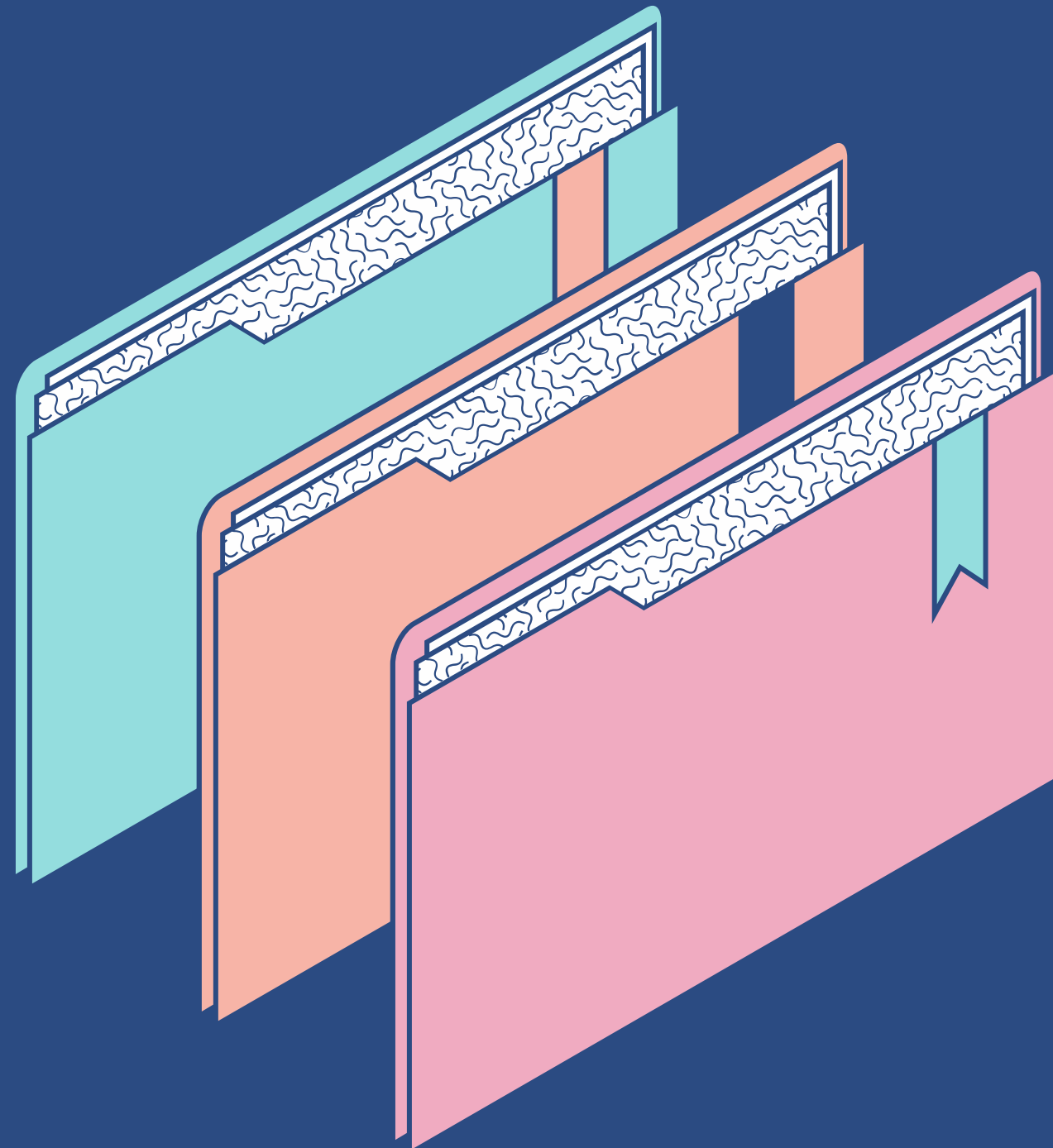
# Comparison Algorithm KNN, SVM, Decision Tree dan MLP in disease classification HEPATITIS C

Member:

Marsya Jelita

Rifqi Alfinnur Charisma

Titan Bagus Bramantyo

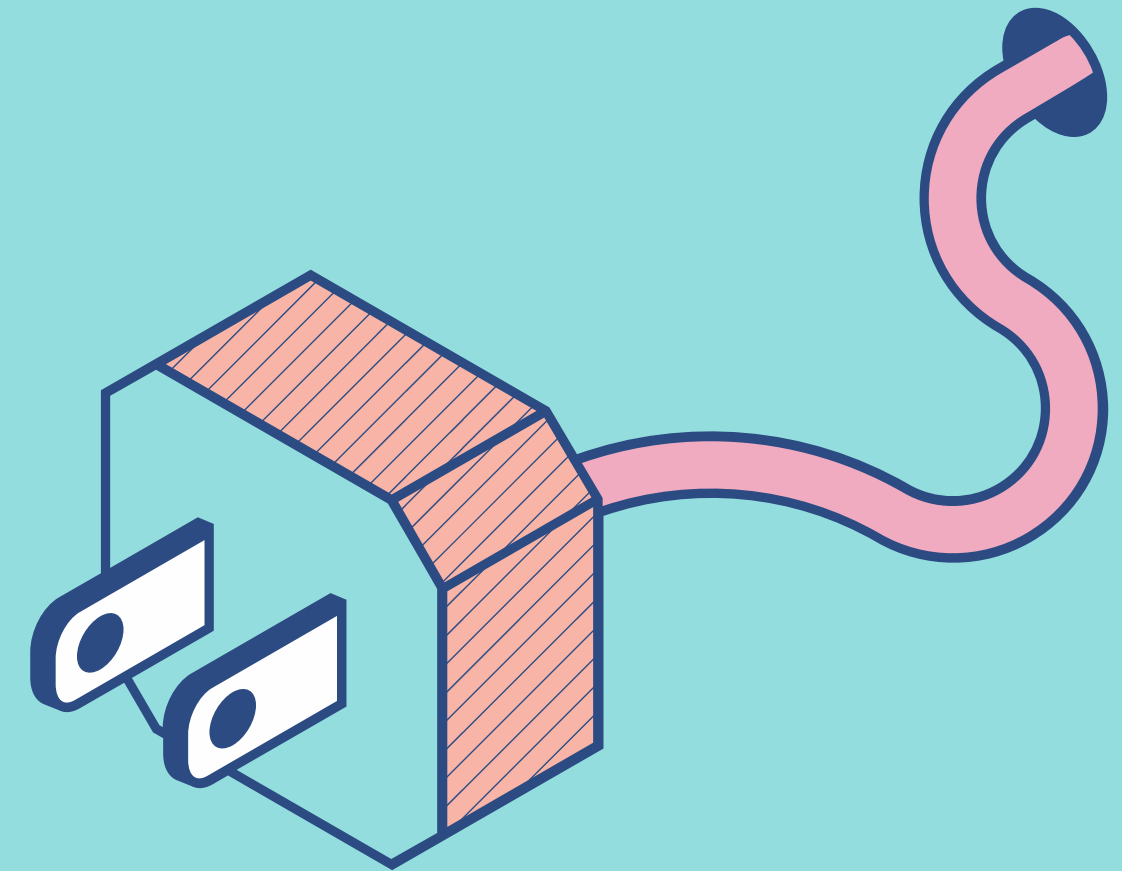


# Table Of Content

- **Background**  
brief presentation on the topic of analysis
- **Data Processing Flow**  
explain the data preparation
- **Model Evaluation Metrics**  
algorithm comparison
- **Conclusion**  
the answer of analysis objective points

# Background

Hepatitis virus infection is an infection that attack the liver. Hepatitis C virus causes acute and chronic infection. Acute HCV infection has no symptoms and almost not cause disease. Along with the development of technology, the presence of artificial intelligence has attracted a lot of attention in the health sector, especially for predicting certain diseases such as Hepatitis C.

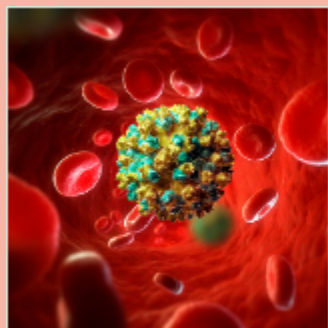


# About Dataset

To accommodate this case analysis process, we use the following data.

Our dataset is taken from Kaggle.com which contains data from laboratory values of blood donors and Hepatitis C patients and demographic values like age.

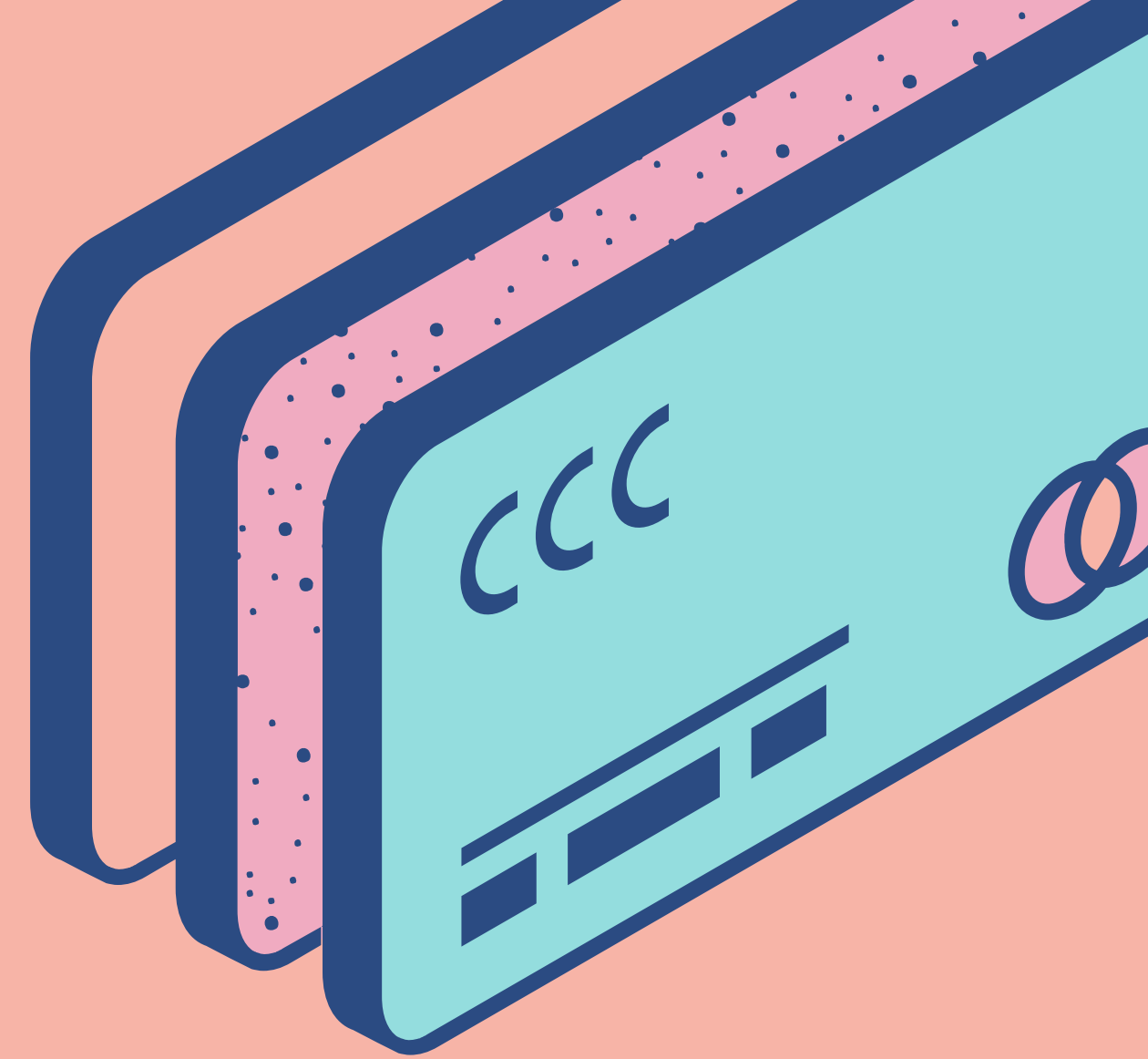
The data contains 615 observations and 14 attributes



## Hepatitis C Prediction Dataset

Laboratory values of blood donors and Hepatitis C patients

[k kaggle.com](https://www.kaggle.com)

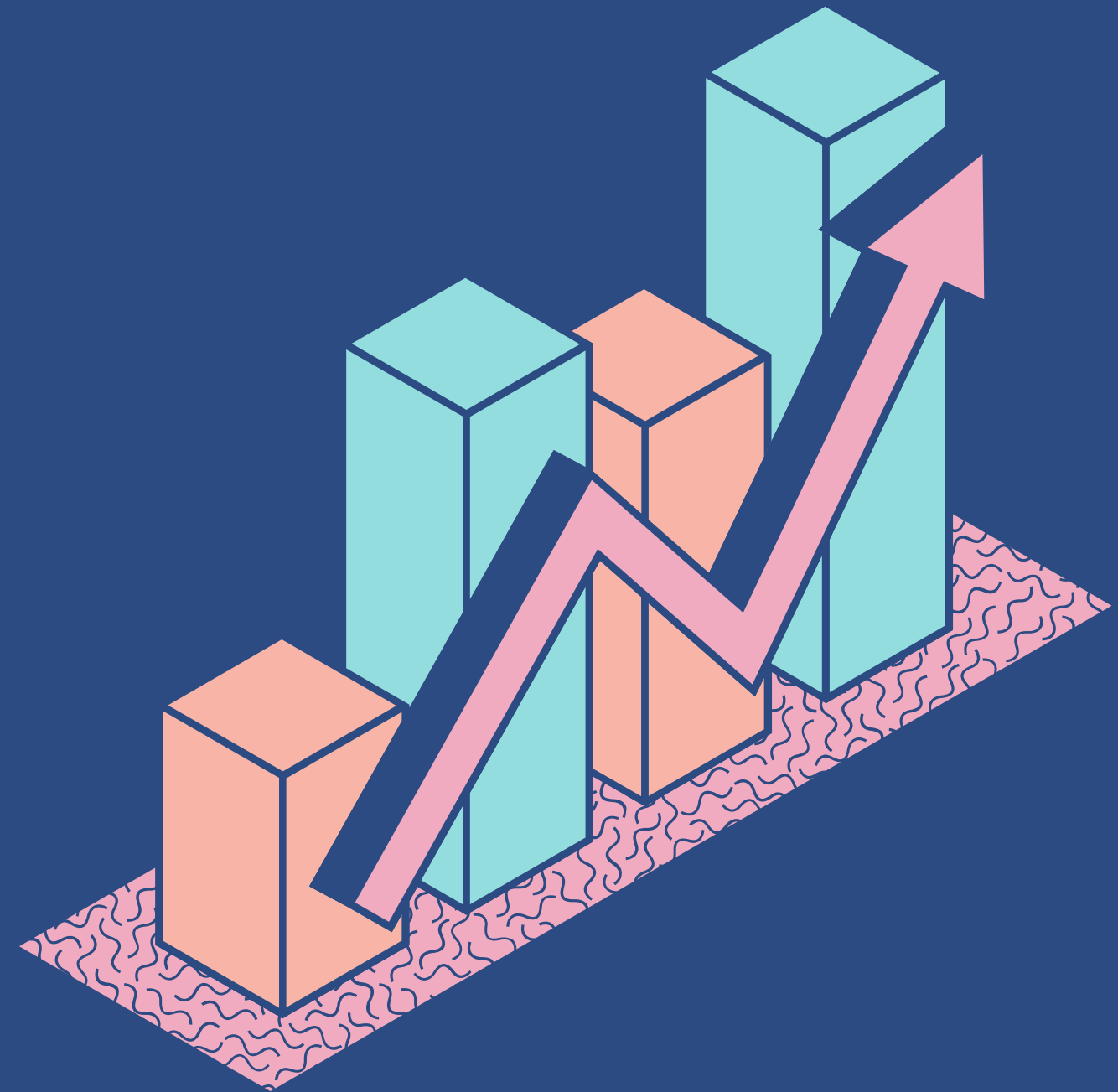


# Features

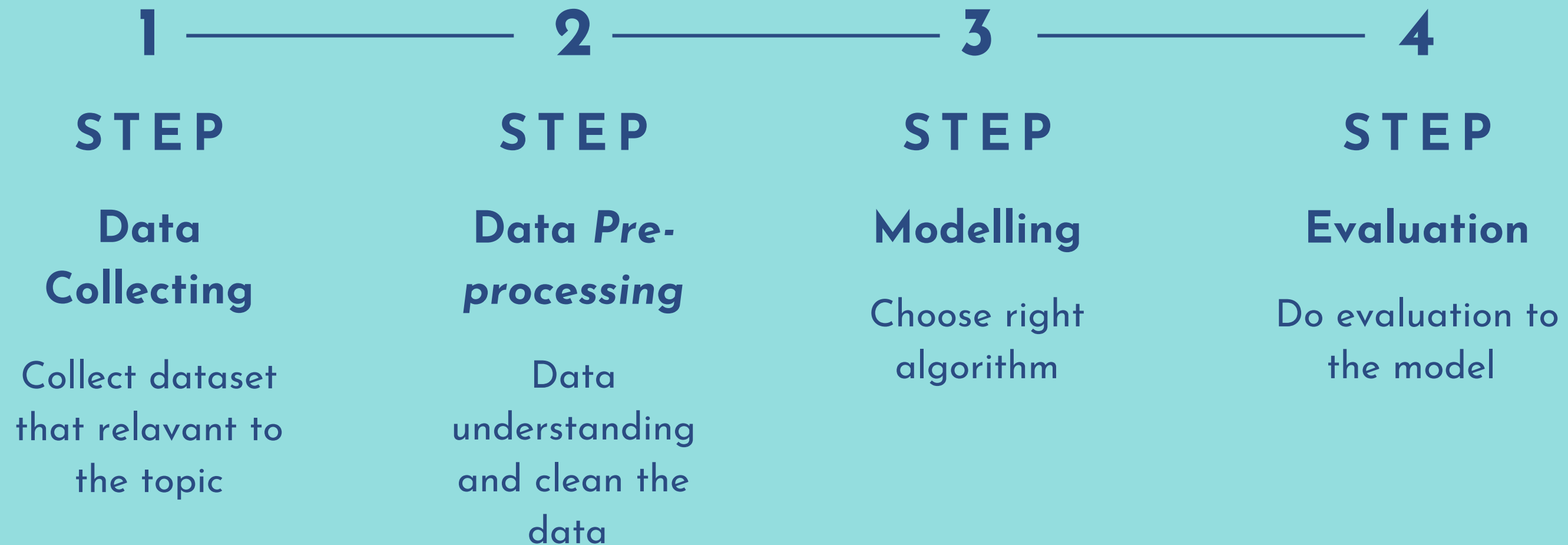
- Category (diagnosis) (values: '0=Blood Donor', '0s=suspect Blood Donor', '1=Hepatitis', '2=Fibrosis', '3=Cirrhosis')
- Age (in years)
- Sex (f,m)

Attributes 5 to 14 refer to laboratory data:

- ALB
- ALP
- ALT
- AST
- BIL
- CHE
- CHOL
- CREA
- GGT
- PROT



# Data Processing Flow



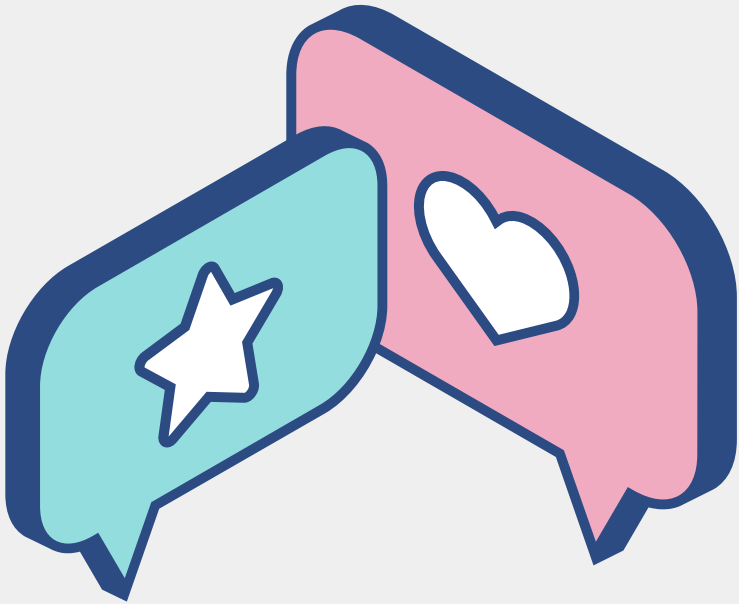


# Algorithm that we use:

- SVM
- KNN
- DECISION TREE
- MLP

# Model Evaluation Metrics

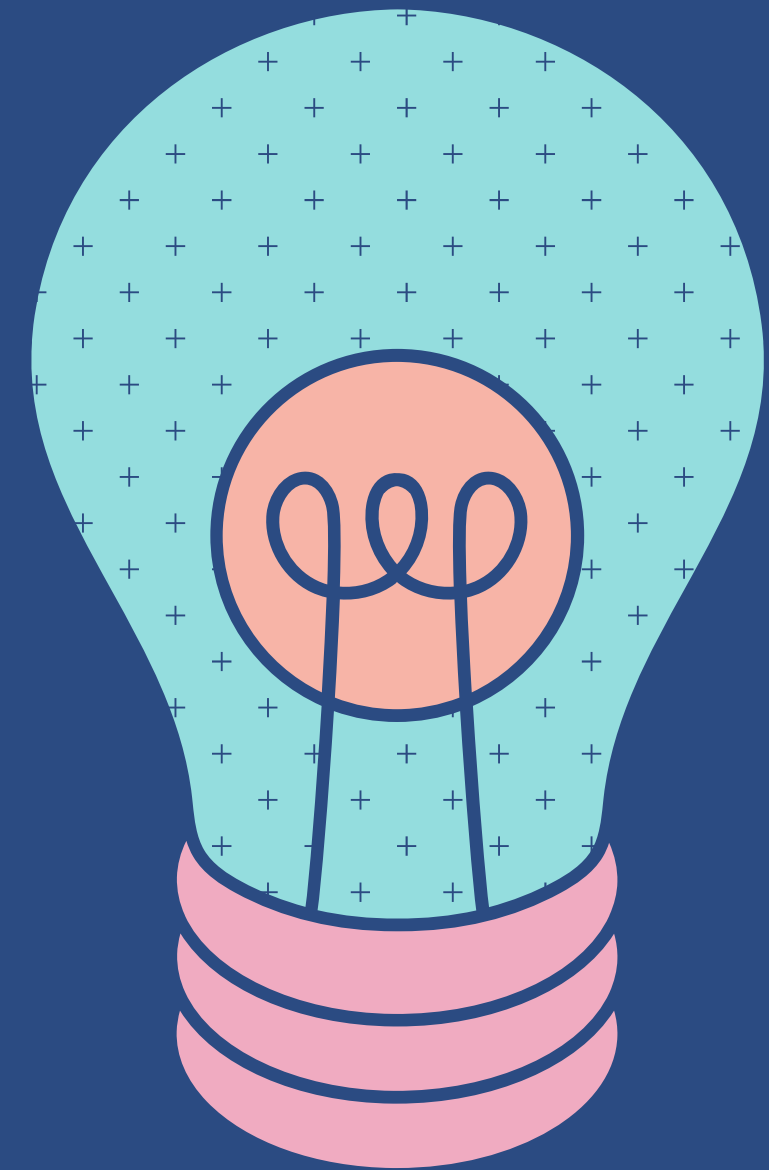
Model	Accuracy	Precision	Recall	F1-Score
SVM	98,17%	88%	82%	84,5%
KNN	86,99%	93%	66,5%	71,5%
DECISION TREE	92,68%	92%	86,5%	89%
MLP	95,12%	97%	87,5%	91,5%





# Conclusion

Based on the four algorithm that we use for build the model, the final result shows that MLP algorithm have the highest score in precision, recall and f1-score. While, highest score in accuracy is from SVM algorithm.





# **Why MLP have Highest score in most evaluation matrix?**

Because MLP is a deep learning algorithm so it has deeper learning model than machine learning and MLP can fix data non linier problems.

# Arigatou!

