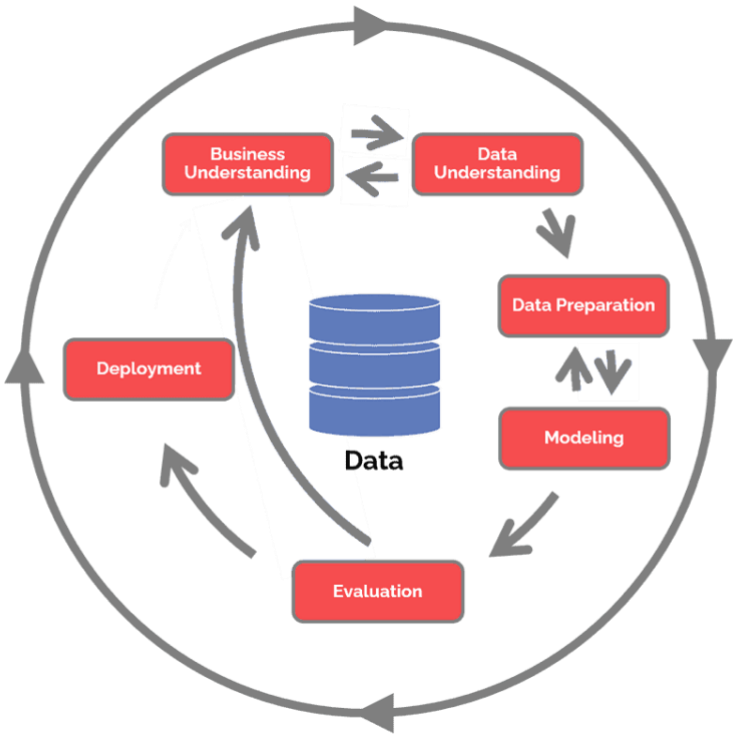
Project Title : BioAct: AI-Driven Drug Activity Classification				
Project Start Date : 12 - 02 - 24		**Project End Date** : 27- 04 -24		**No.of Months** : 04
Sl.No	**VTU No**	**Name of the Student**	**Branch**	**Year of study**
1.	**19283**	**G. Lokesh Reddy**	**CSE**	**3RD**
2.	**19234**	**I. Dineesh**	**CSE**	**3RD**
3.	**19235**	**T. Ganesh**	**CSE**	**3RD**

Abstract :

Predicting a drug's Mechanism of Action (MoA) is crucial for efficient drug discovery. This project aimed to develop a machine learning model to predict MoA for unseen drugs based on unique dataset that combines gene expression and cell viability data. The first major step performed is Exploratory Data Analysis (EDA) to understand the gene expression, cell viability data, and drug labels. Then experimented with various machine learning models commonly used for multi-label classification, including Logistic Regression, Random Forest Classifier, Gradient Boosting Classifier, and GaussianNB. Additionally, XGBoost and LightGBM were also used for comparision. Then the evaluation of the models was done on the basis of classification report metrics, log loss and employed hyperparameter tuning to improve their performance. The best performing model achieved was Gradient Boosting Classifier. This project demonstrates the potential of machine learning approaches for MoA prediction, contributing to advancements in drug discovery.

Block Diagram :



Project Outcome/ Result:

- **Development of a Machine Learning Model for MoA Prediction:** The core outcome is a machine learning model that can analyze cellular response data (gene expression and cell viability) and predict the MoA for unseen drugs. This model should improve the accuracy of MoA prediction compared to traditional methods.

- **Improved understanding of the relationship between cellular responses and MoA:** By analyzing the data and building the model, you'll gain insights into how gene expression and cell viability patterns are linked to specific drug mechanisms.

- **Contribution to advancements in drug discovery:** A successful MoA prediction model can significantly accelerate drug development by allowing scientists to focus on targeted therapies with a higher likelihood of success.

Conclusion:

This project investigated the application of machine learning to predict a drug's Mechanism of Action (MoA). Leveraging several machine learning models suited for multi-label classification, including Logistic Regression, Random Forest Classifier, Gradient Boosting Classifier, and GaussianNB, the project explored the relationships between gene expression, cell viability data, and MoA labels through Exploratory Data Analysis (EDA). Classification report metrics were used to evaluate model performance, and hyperparameter tuning was employed to optimize model accuracy.This project underscores the potential of machine learning approaches for MoA prediction based on cellular response data. The developed model offers a valuable contribution to advancements in drug discovery by empowering scientists to prioritize targeted therapies with a greater likelihood of success. Future endeavors could explore more complex models, integrate additional biological information, and further refine the prediction accuracy for MoA classes of particular interest.

List of Algorithms Considered	Name of your Evaluation Metric	What is the Evaluation Score
Logistic Regression	Accuracy	68
Random Forest	Accuracy	68
Gradient Boosting Classifier	Accuracy	70
GaussianNB	Accuracy	46
XGBoost	Accuracy	67
LightGBM	Accuracy	68

Which Algorithm is Finalized : Gradient Boosting Classifier