



**Pragati**  
**Engineering**  
**College**  
An Autonomous Institution  
Accredited by NBA & NAAC with 'A' grade

## DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

# SHARK TANK - STARTUP SUCCESS AND FAILURE PREDICTION USING MACHINE LEARNING ALGORITHMS

20A31A4205	C.S.S.Swetha
20A31A4208	D.Sreeja
20A31A4214	K.Naga Sri Lakshmi
20A31A4206	S.Charishma
20A31A4227	U.Nikitha

### Project Supervisor

Mrs. P. Satyavathi (Asst. Professor)  
Dept. of CSE(AI & ML)

### Head of the Department

Dr. A. Radha Krishna  
Dept. of CSE(AI & ML)

# Abstract

Startups are the growth indicators of a country. Growing startups are a good opportunity for investors to get high returns in less time and less capital. In the competitive world of startups, predicting success is crucial for strategic decision-making. With 90% of startups facing challenges, employing robust methods to forecast success is essential. This study leverages Machine Learning algorithms like AdaBoost, Gradient Boosting, and Random Forest to evaluate key performance indicators. AdaBoost stands out with a predictive accuracy of 78%, highlighting its effectiveness. Using extensive historical data, our models offer insights into the factors influencing startup success, guiding stakeholders and investors in making informed decisions.

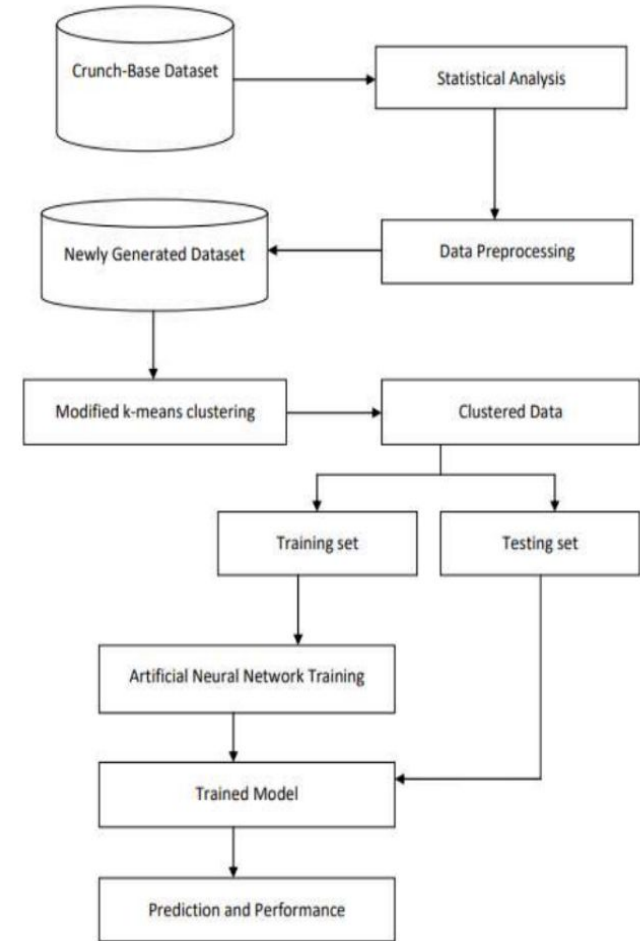
# Introduction

- Shark Tank is a platform where startups try to pitch their business ideas to a panel of investors in hopes of securing funding
- Manually analysing startups & making decisions is time-taking & complex process
- ML algorithms predict startup success by analyzing factors like market demand, financial metrics, etc. and promise investors and entrepreneurs to make informed investment decisions
- Leveraging machine learning could increase the chances of success for new businesses while reducing the risk of failure



# Existing System

- Predicts operational status of the company ["Operating", "Acquired", "Closed", & "IPO"]
- First, modified k-means clustering is used to pre-train the model and group data according to feature similarity
- Then, grouped data is being used with the ANN for training and prediction of the startup's success
- Limitations: Cannot handle noisy data, less generalizability and accuracy

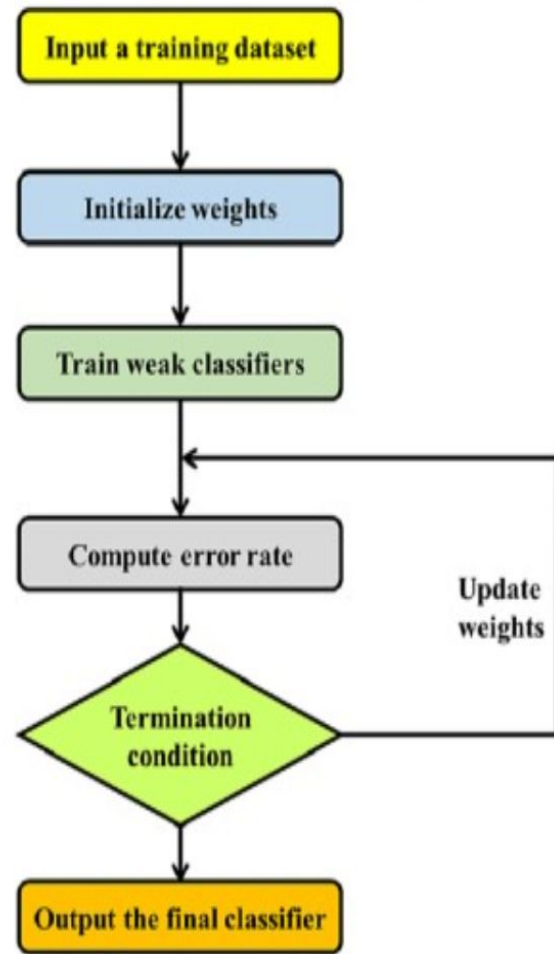


# Proposed System

- Predicts operational status of the company as [“Success”, “Fail”]
- Compared and Contrasted three algorithms - AdaBoost, Gradient Boost and Random Forest to analyze and evaluate startup data
- Model Evaluation: Metrics such as Accuracy, F1 Score, Precision & Recall are considered
- Many factors such as type of Industry, no. of Competitors, Funding total, Funding Rounds, Venture Capitalists, Angel Investors are analysed

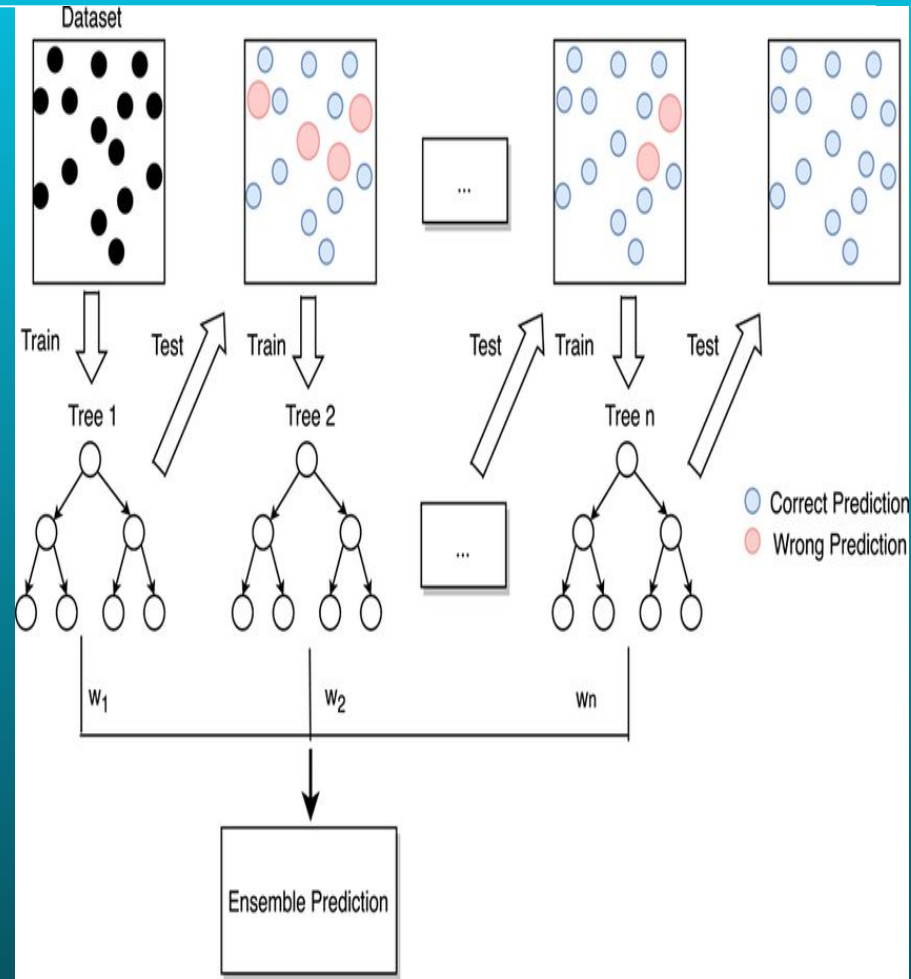
# AdaBoost Algorithm

1. **INITIALIZE** - Assign equal weights to all training samples
2. **ITERATION** - Until error rate is minimized
  - Train a weak learner(decision tree) on weighted dataset
  - Calculate the error rate and update weights of misclassified points
  - Compute classifier weight
3. **MODEL** - Combine the weak learners using their computed weight to form a strong classifier
4. **PREDICTION** - For a new sample, prediction is determined by weighted majority vote of all the weak classifiers



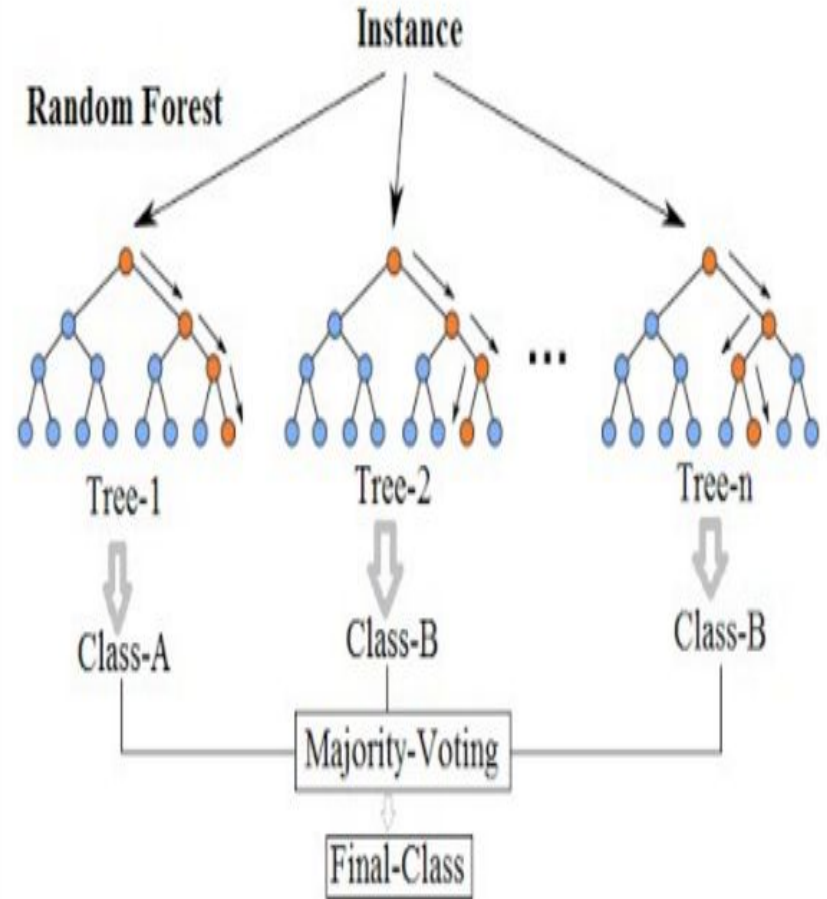
# Gradient Boost Algorithm

1. **INITIAL MODEL** - Predicts the most common class for classification
2. **ITERATION** - Until error rate is minimized
  - Use a weak learner (decision tree)
  - For every data point, calculate residuals
  - **TRAIN NEW MODEL ON RESIDUALS**- If the new model can predict the errors of the previous model, the combined model will be more accurate
  - **UPDATE MODEL** - Previous predictions + New models predictions to improve accuracy weighted by Learning Rate
3. **PREDICTION** - For a new sample, prediction is determined by weighted majority vote of all the weak classifiers



# Random Forest Algorithm

- 1. CREATE BOOTSTRAP SAMPLES** - Randomly pick data points from dataset (with replacement) to create several new datasets
- 2. TRAIN DECISION TREES**
  - Train a decision tree
  - At each split, choose the best split from a random subset of features
- 3. MODEL** - Once all the trees are trained, combine them into one large model called a Random Forest
- 4. PREDICTION** - For a new sample, aggregate the predictions from all the decision trees - majority vote





# Advantages of Proposed System

- **Enhanced Predictive Accuracy:** Utilizes AdaBoost, Gradient Boosting, and Random Forest algorithms; with AdaBoost achieving 78% accuracy
- **Mitigation of Overfitting:** Ensemble methods effectively handle complex data relationships and reduce overfitting
- **Clear and Actionable Outputs:** Outputs either “Success” or “Failure” which is simple to understand and act upon

# Tech Stack

Technology Required	
Programming Languages	Python
Back-End Frameworks	Django
ML Algorithms Used	AdaBoost Algorithm, Gradient Boost Algorithm, Random Forest Algorithm
Front-End Technology	HTML, CSS, JavaScript
Database Language	SQL
Collaboration	VS Code

# Results



# Input & Output Formats

Investors :

Average Participants :

Competitors :

Founders :

Top 500 :

Yes

▼

Funding rounds :

▼

Funding-Total:

Relationships :

Which Industry ?

☐ Software

☒ Enterprise

☐ Web

☐ Mobile

☐ Advertising

☐ Games-Video

☐ E-commerce

☐ Biotech

☐ Consulting

☐ Other-Catagory

Has Venture Capitalists :

Yes

▼

Has Angel Investors:

Yes

▼

Start Prediction

## PREDICTION RESULT

Our prediction says that your startup will hopefully be a Success. Forget about your competitors just focus on your customers.



## PREDICTION RESULT

Our prediction says that your startup might fail. As funding rounds are less, try to increase funding rounds then you will definitely succeed.



# Conclusion & Future Enhancements

- Investors are highly interested in making money by investing in startups as they are able to deliver high returns in less amount of time
- Therefore, analysis of success and failure of startup is beneficial for investors, company administrators, and governmental agencies
- Several algorithms are in place to automate predictive analysis of startups
- Our AdaBoost algorithm proved to show better results compared to different earlier versions
- Future enhancements include incorporating customer reviews for more accurate predictions and integrating unstructured data sources