



STARTUP SUCCESS PREDICTION USING MACHINE LEARNING

A data-driven approach to predict startup outcomes

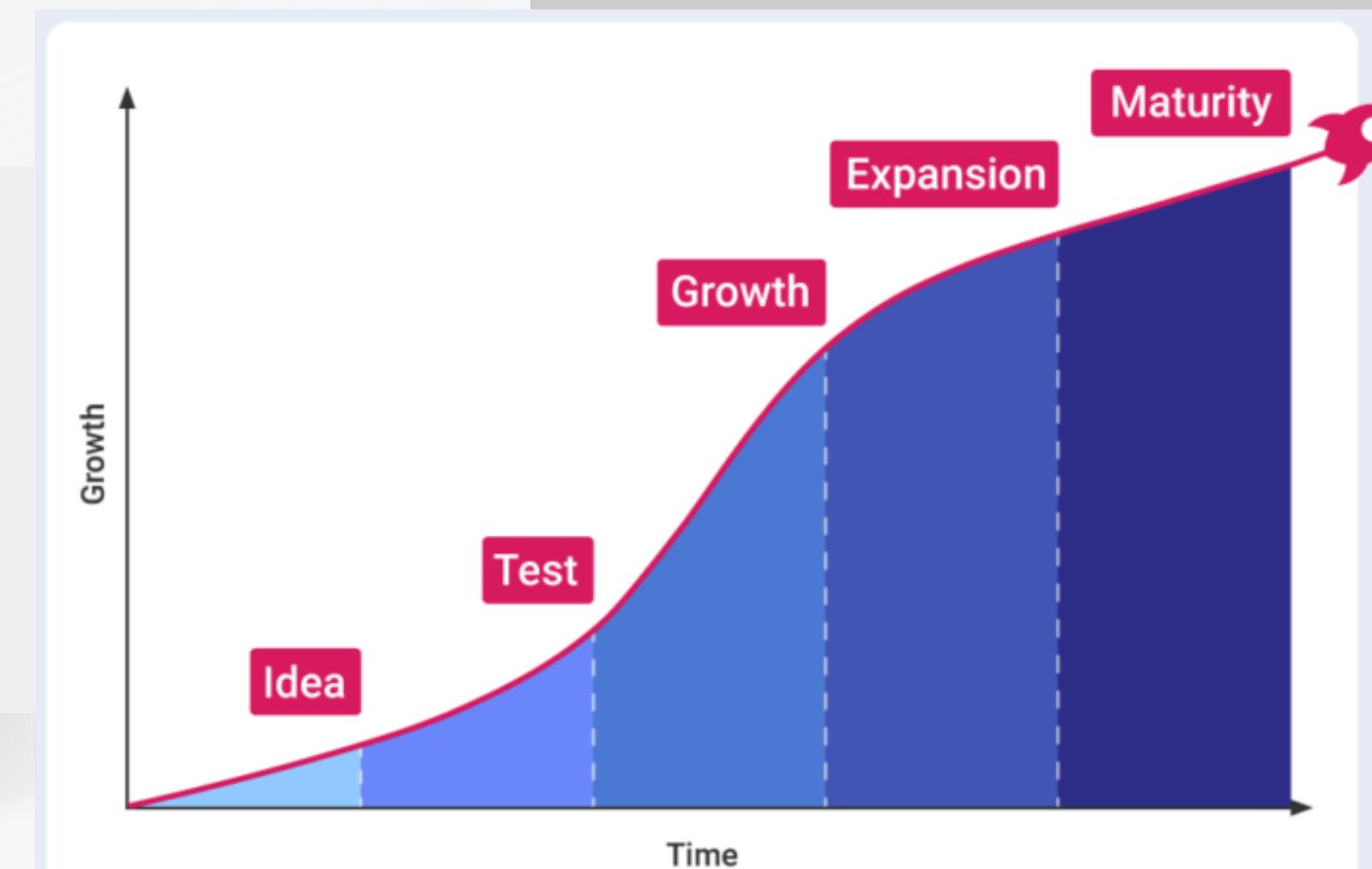
OUR TEAM

- Belal Ibrahim Fareed
- Mahmoud Mahmoud Ali
- Mahmoud Mohamed Elshirbeny
- Mohamed Khaled Ibrahim
- Sara Adel Shokry

WHY PREDICTING STARTUP SUCCESS MATTERS



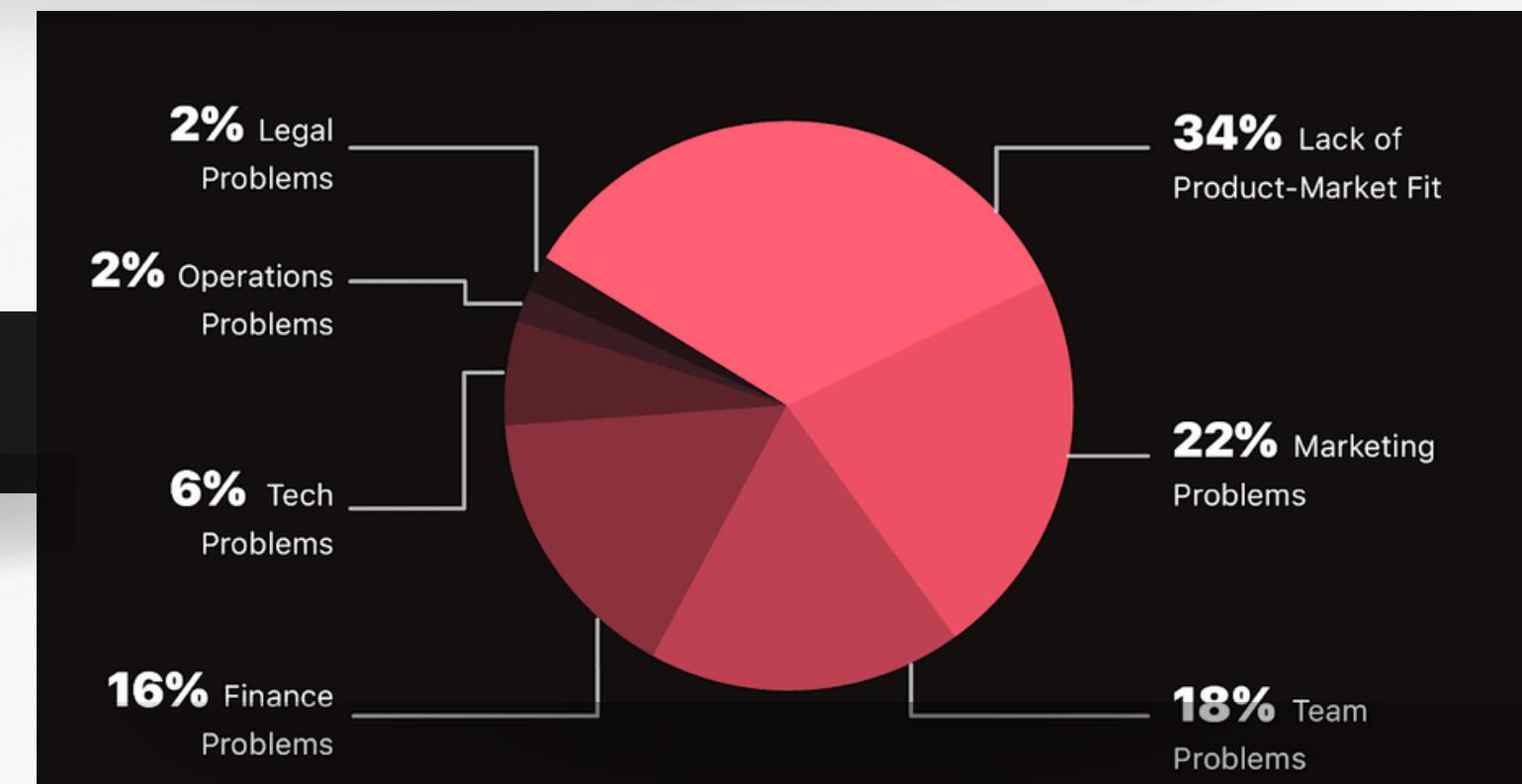
- Startups are essential for innovation and economic growth.
- They face high uncertainty, with a majority failing early.
- Early predictions can help investors identify promising startups.



CHALLENGES IN THE STARTUP ECOSYSTEM

- High failure rates and uncertain outcomes.
- Need for a scalable and data-driven method to assess potential success.
- Helps investors and entrepreneurs make informed decisions.

Vision



PROJECT OBJECTIVE

Objective n° 1

Predict whether a startup will succeed (acquired) or fail (closed).

Objective n° 2

Use machine learning to analyze historical startup data.

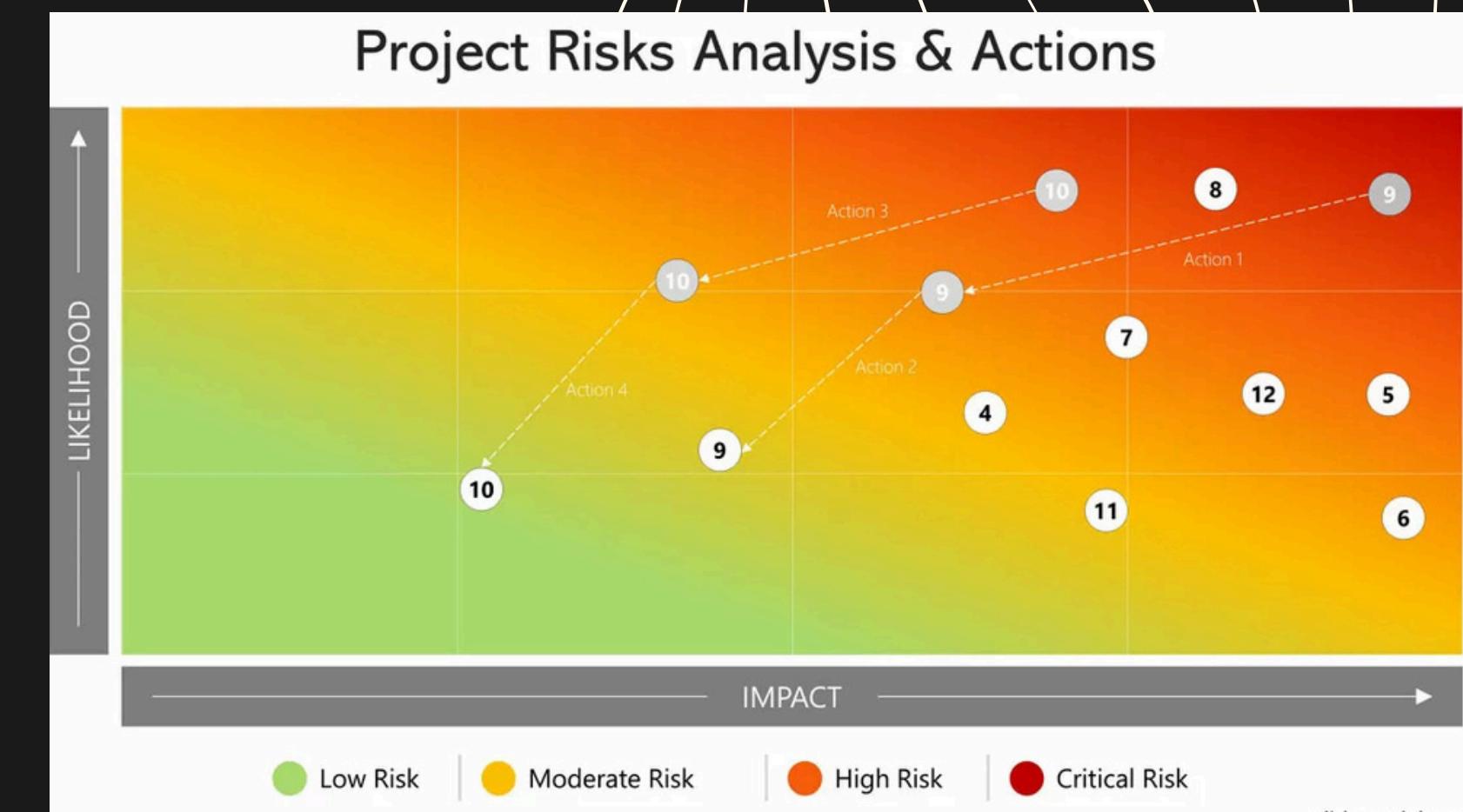
Objective n° 3

Provide actionable insights for stakeholders.



Dataset Overview

- Data from [Dataset Source, if any].
- 48 features, including funding details, milestones, relationships, etc.
- Target variable: status (acquired/closed).



Transforming the Data

Before cleaning and enriching

Doe John C., PhD

Name, middle
name, surname

Director IT Applications
at ITeam | Back-end

Employed
Member
description

Job title
Employee
department

Employee
subdepartment

Employee
management level

ITeam;

Lebanon

<p class="description" data-section="currentPositions">Facilitate networking and cultural navigation</p>

752, 696, 395

After cleaning and enriching

John C. Doe

Yes

Director IT Applications
at ITeam | Back-end

Director IT Applications

Information Technology

IT Operations

Director

ITeam

Lebanon, New Jersey,
United States

Facilitate networking
and cultural navigation

personal development,
quality assurance,
entrepreneurship

01

02

03

04

Removed irrelevant
columns like name,
id, etc.

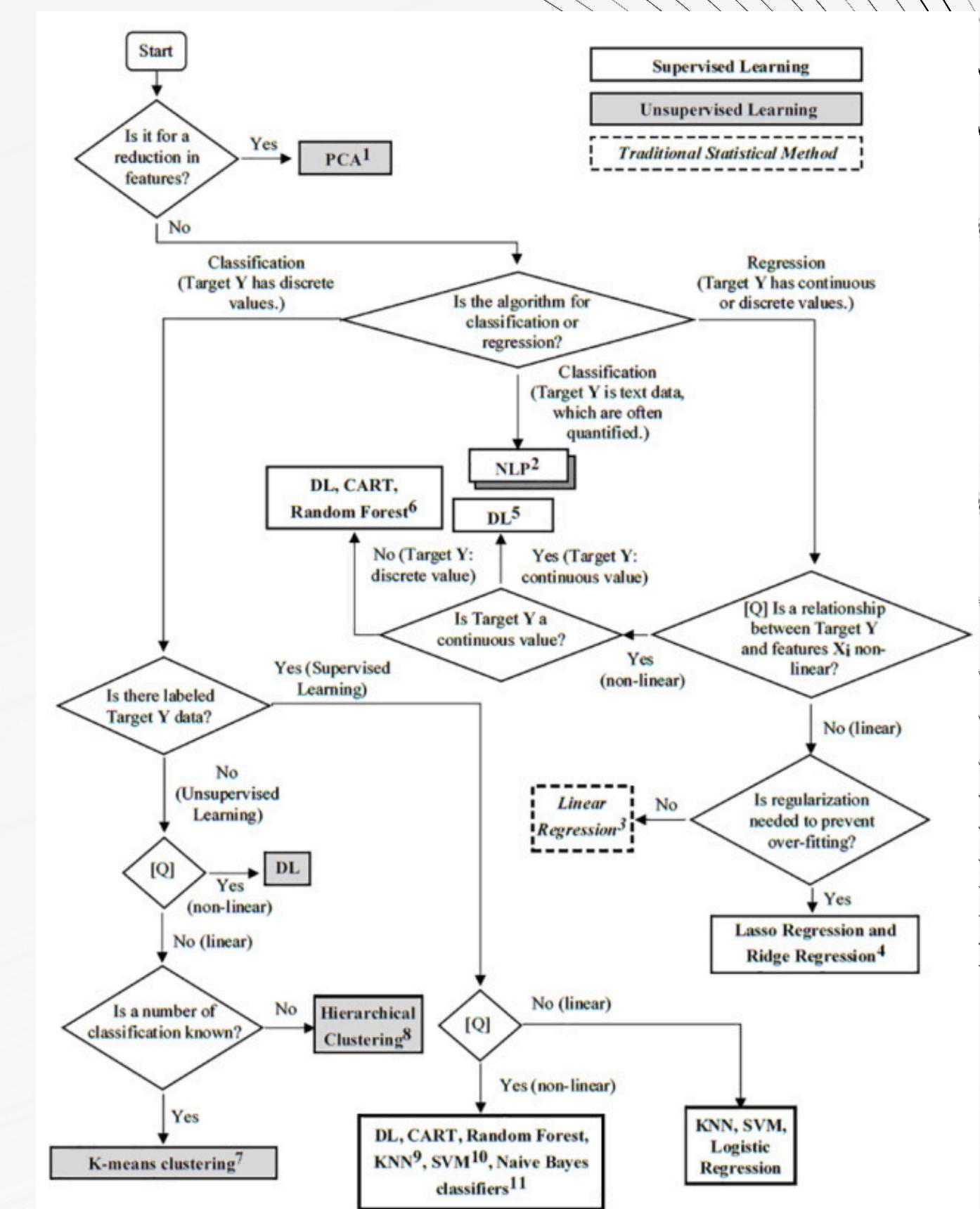
Handled outliers in
avg_participants.

Encoded categorical
variables (e.g., status
to label).

Split data into
features (X) and
target (y).

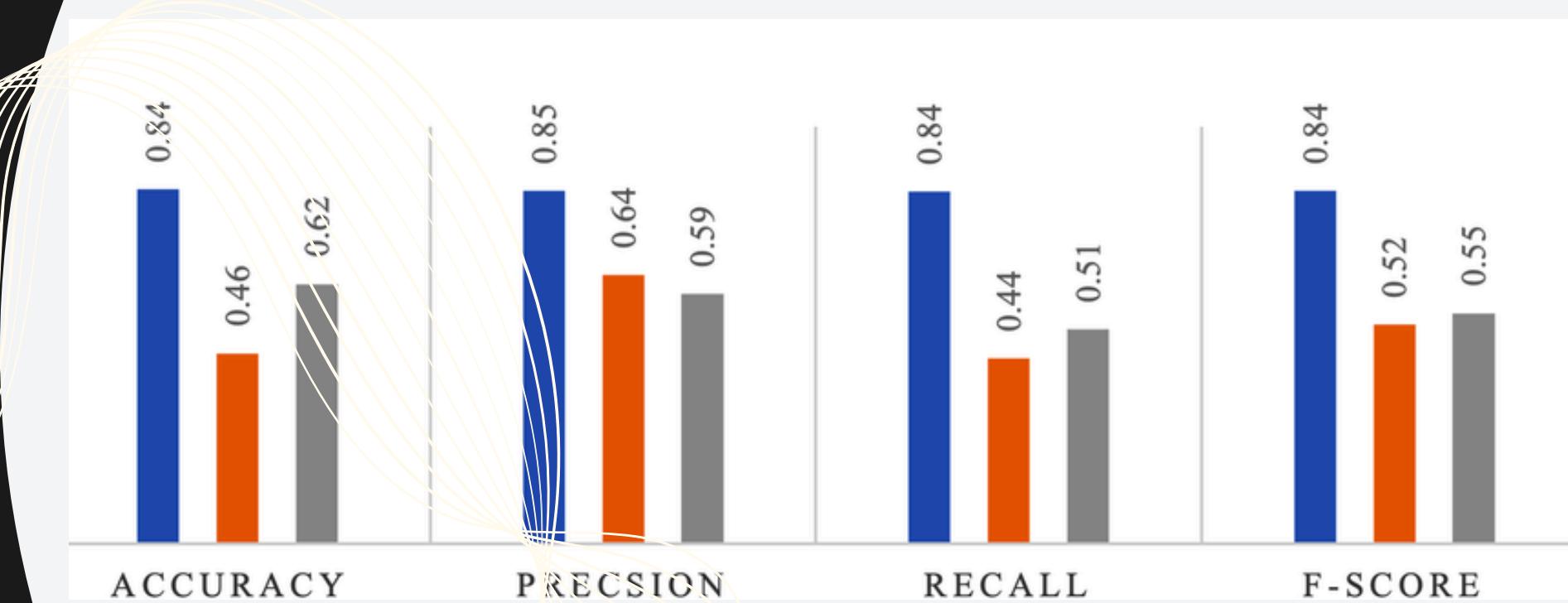
CHOOSING THE RIGHT MODELS

- Logistic Regression: Simple, interpretable baseline.
- Decision Tree Classifier: Captures complex relationships in data.
- Used supervised learning as we have labeled data (status).



MODEL PERFORMANCE

- Logistic Regression:
 - Accuracy: 68.64%
 - F1-Score: 43.90%
- Decision Tree:
 - Accuracy: 100%
 - F1-Score: 100%
- Single Example Prediction:
 - Failure (Closed)



WHAT DID WE LEARN?

Logistic Regression is less complex but has moderate accuracy.

Decision Tree perfectly fits training data (risk of overfitting).

Features like funding_total_usd, milestones, and relationships are critical.

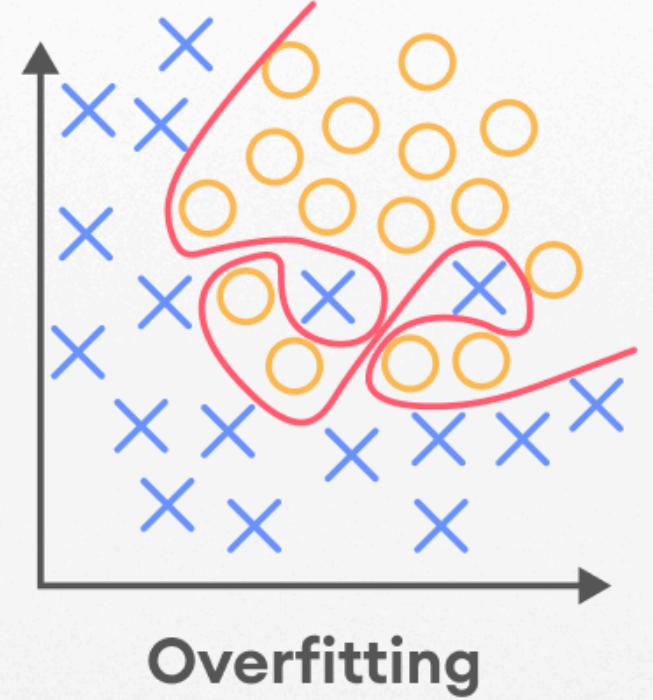
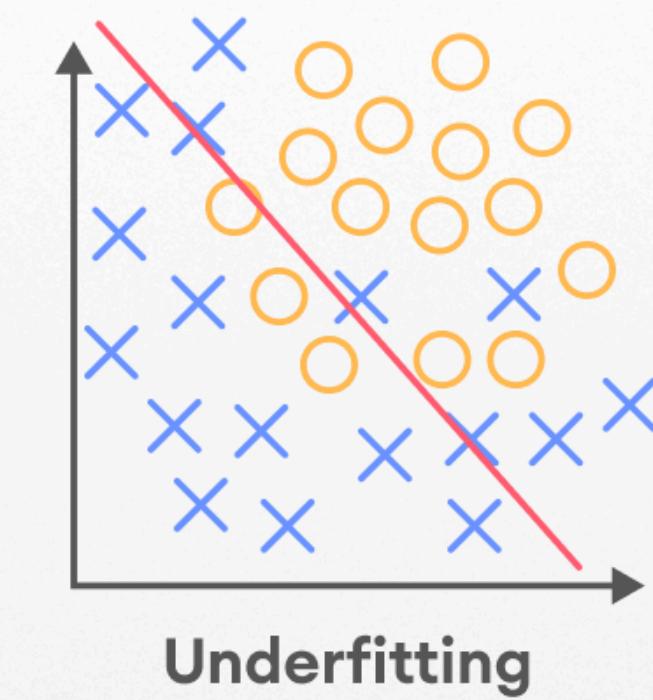
REAL-WORLD USE CASES

- Investors: Identify high-potential startups.
- Startup Founders: Understand key success factors.
- Accelerators/Incubators: Filter applicants for programs.



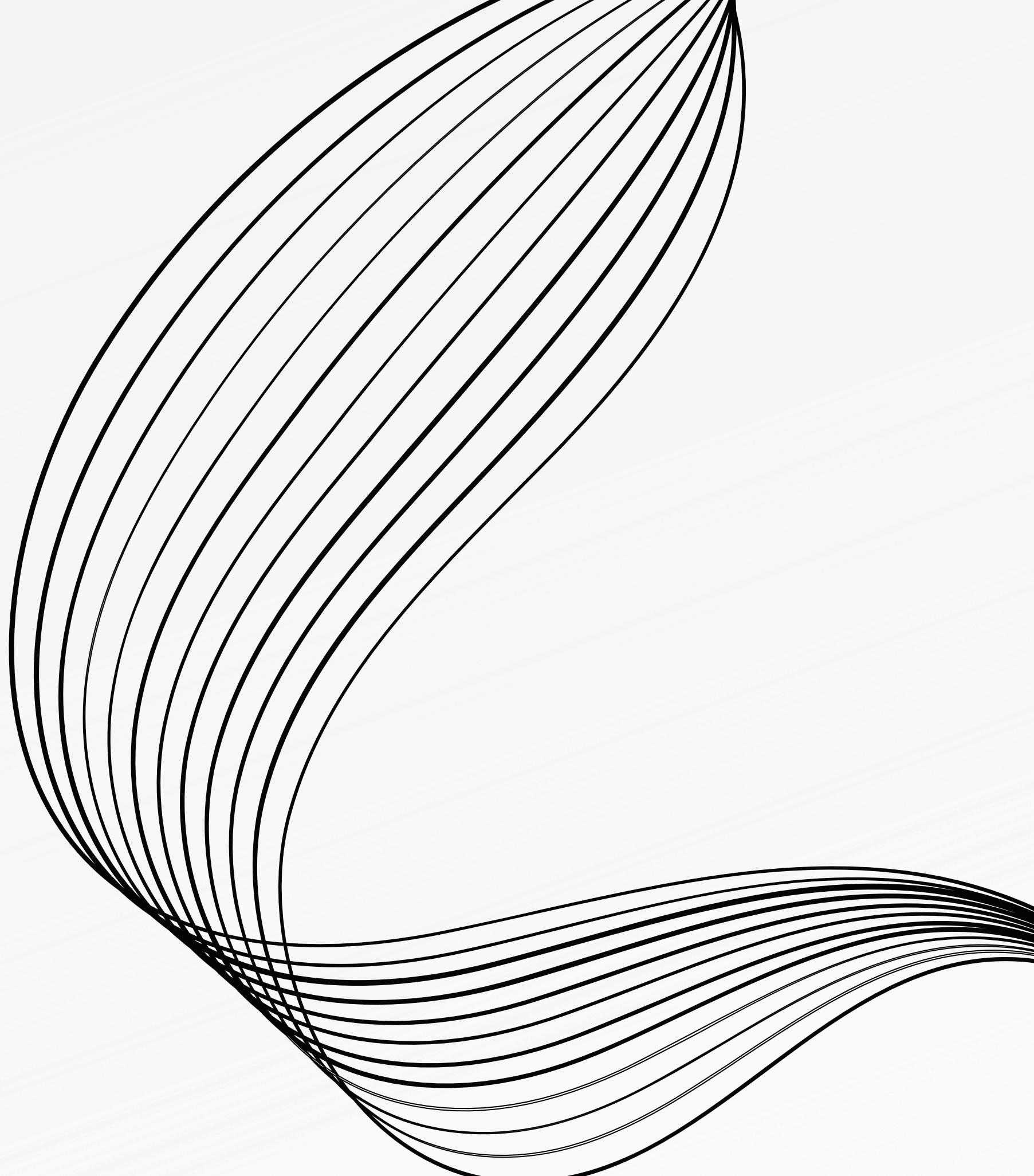
LIMITATIONS AND NEXT STEPS

- Challenges:
 - Imbalance in success vs. failure data.
 - Decision Tree overfitting.
- Future Work:
 - Use more diverse data sources.
 - Experiment with advanced models (e.g., Random Forest, XGBoost).
 - Address overfitting using cross-validation and pruning.



FINAL THOUGHTS

- *Machine learning offers a valuable tool for predicting startup success.*
- *Insights from the model can guide stakeholders in decision-making.*
- *Future enhancements can improve accuracy and robustness.*



RELATED WORKS

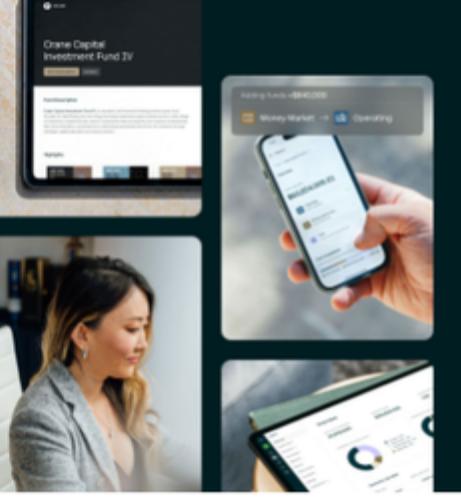
Built to scale all private funds

AngelList

Build, Lead, Invest

AngelList builds the infrastructure that powers the startup economy—providing investors and innovators with the tools to grow.

 AngelList



Private market investing made easy

RAISE FUND. START SYNDICATE. DRIVE GROWTH

For investors For startups

 letsventure

 trica CAPITAL

 trica EQUITY

Boosting India's private markets

Empowering Indian startups, investors, and family offices. Streamlining capital and growth for Indian startups

 letsventure.com

<https://www.sciencedirect.com/science/article/pii/S0040162523002998>

THANK YOU