

# BizLens: Startup Success Evaluator

We built an AI-powered tool that predicts early-stage startup success using machine learning and data insights.



# Team Members



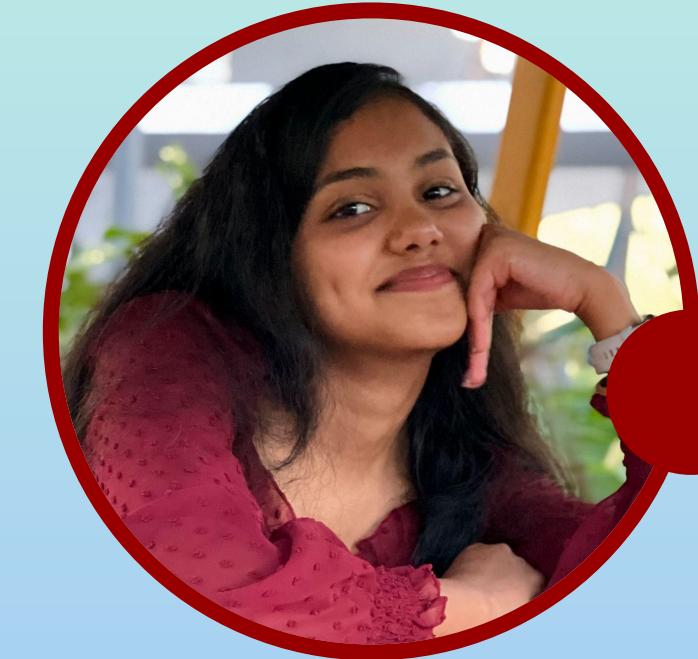
**Ebyan**

University of  
Minnesota  
Computer Science



**Elisa**

Tufts University  
Computer Science  
Cognitive and Brain  
Science



**Shirina**

Florida International  
University  
BA Computer Science



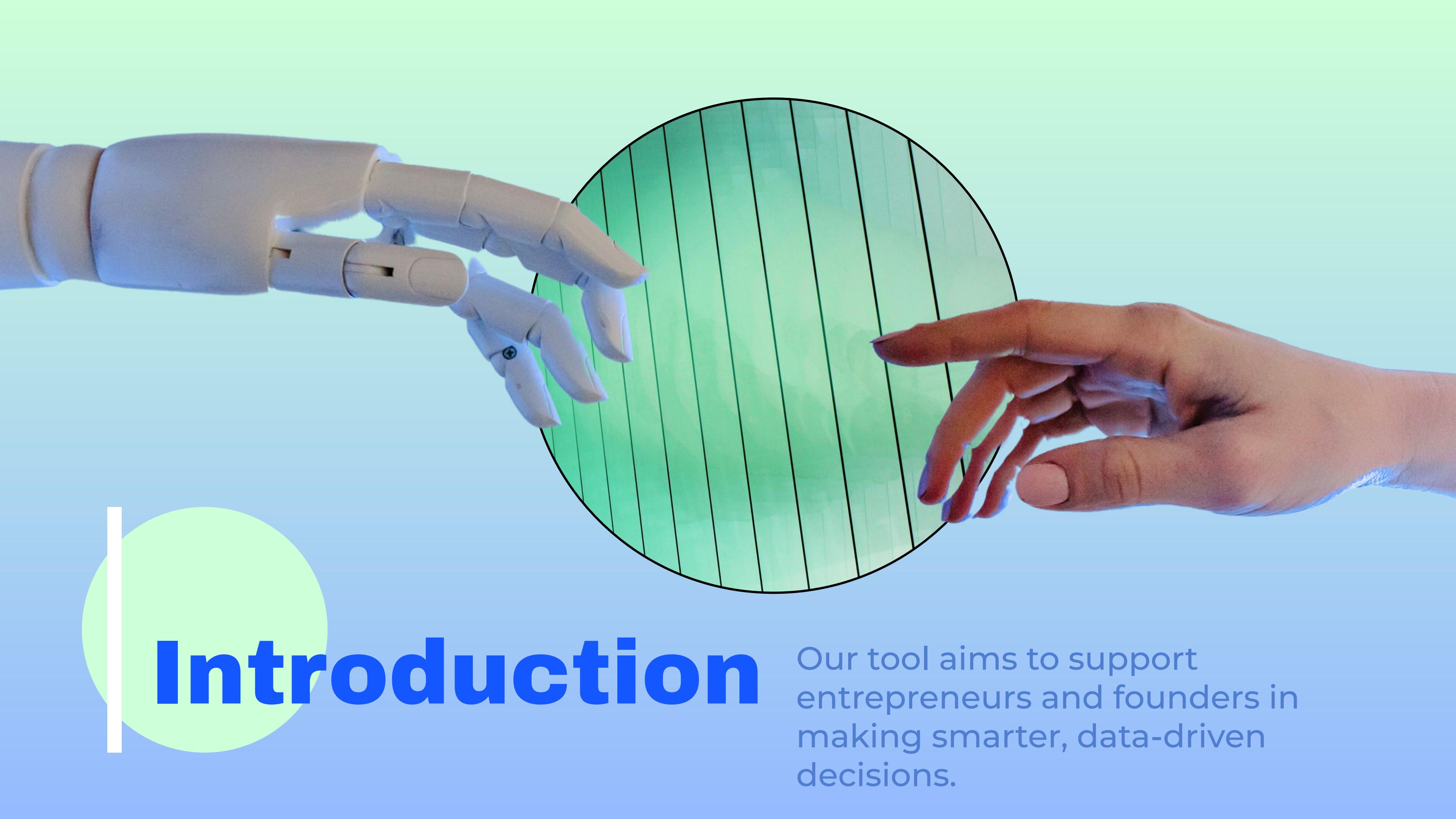
**Sarah**

NYU  
Computer Science  
and Business



**Victor**

Rutgers University  
Computer Science



# Introduction

Our tool aims to support entrepreneurs and founders in making smarter, data-driven decisions.

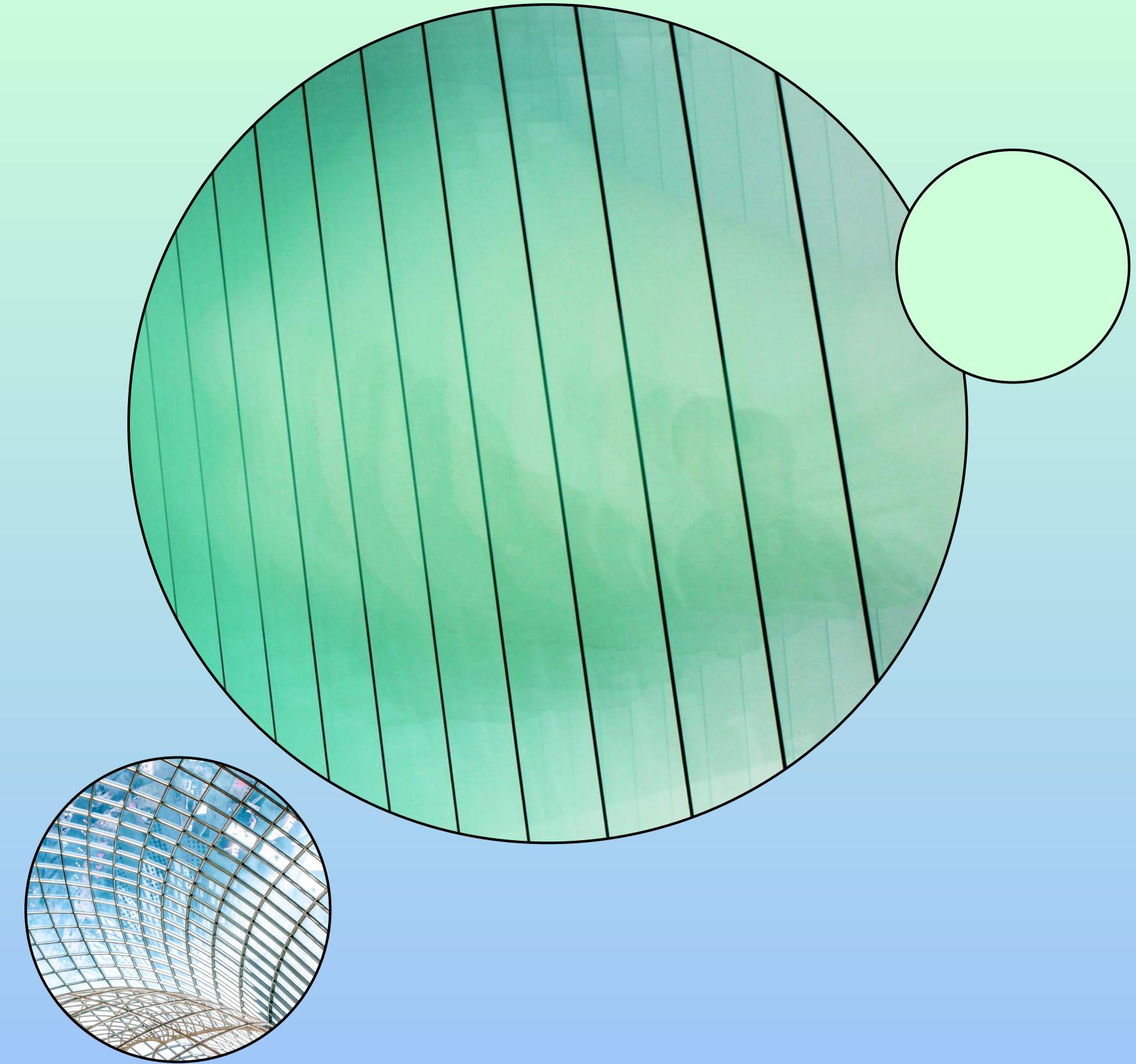
# Topic Summary

For many entrepreneurs, starting a company is a leap of faith. Many have ideas for potential startups, but don't have the business expertise to know what makes a startup successful. With BizLens, we aim to guide entrepreneurs in the early stages of their business by training an AI model to analyze patterns in funding, industry, and team dynamics to determine the startup's chances for success.



# **Research Question**

What features are best predictors of startup success rates, and what actionable insights can we provide to improve their chances of success?



# Project Evolution

## Initial Idea

Build a classifier to predict if a startup will succeed

## Pivoted to Regression

Switched from classification to regression to output a success score

## Exploring Datasets

Cleaned and analyzed global startup dataset

## Thought about UI design

Designed a front-end UI ensuring clarity and user experience

## Trained & Tested Models

Experimented with 3 different algorithms

## Discussed Challenges

Limited data, noisy labels, and potential real-world biases



**Scan to view our  
Github Pages**

The background features a large, light blue circle on the right side. Overlaid on it is a smaller, light green circle on the left. Both circles have thin black outlines. Within the light green circle, there are several thin, dark green vertical lines of varying lengths, creating a striped effect.

# Walkthrough

BizLens

# How successful is your startup?

Assess your startup's potential with our AI-powered evaluation tool.

Start  
Evaluation

N

## What industry is your startup in?

Technology

Healthcare

Finance

E-commerce

Manufacturing

Education

Real Estate

Retail

Transportation

Entertainment

Back

1 of 4

Next

## Which funding types apply?

(select all that apply)

Seed

Convertible Note

Series A

Series B

Series C

Series D

Series E

Debt Financing

Grant

Product Crowdfunding

Equity Crowdfunding

Private Equity

Post-IPO Equity

Angel

Undisclosed

Venture

Back

2 of 4

Next

**In what year was your startup founded?**

e.g. 2

[Back](#)    3 of 4    [Next](#)

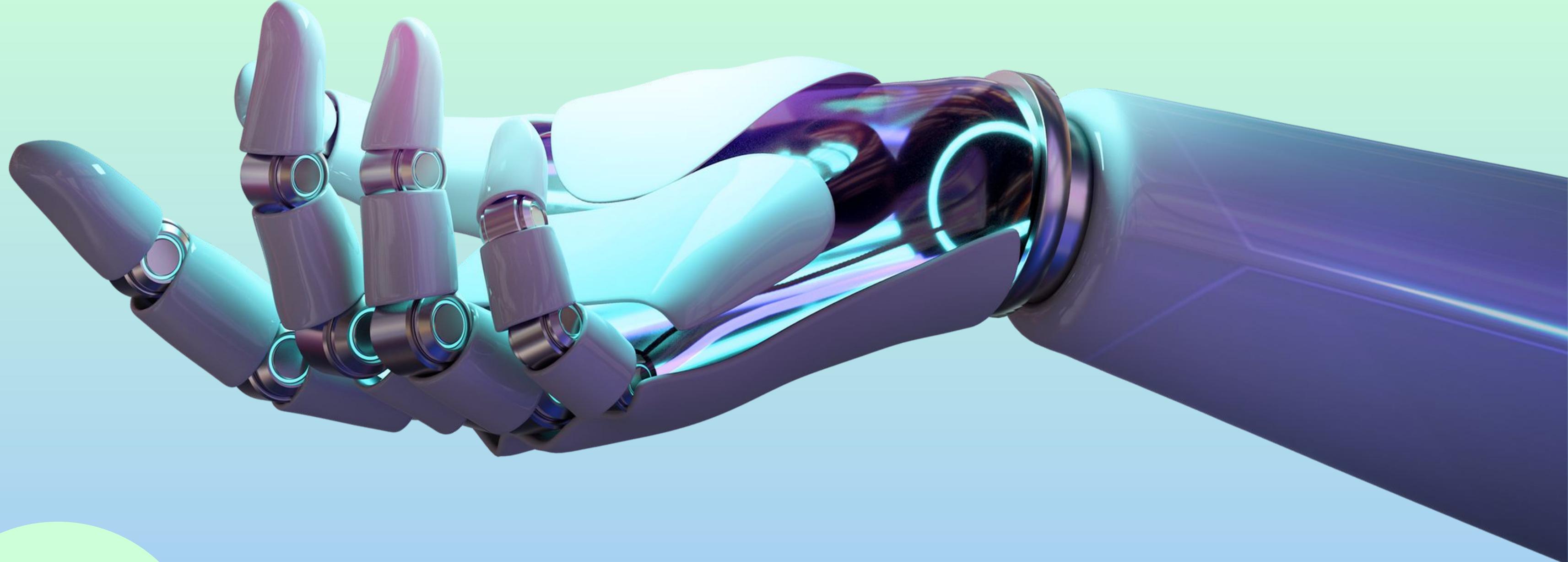
## What is your primary market category?

- ✓ — select one —
- Social Television
  - Enterprise Search
  - Reviews & Recommendations
  - Biomass Power Generation
  - Minerals
  - Parenting
  - Transaction Processing
  - Hardware
  - Auto
  - Health Services Industry

## Your predicted success score



Predict Again



# Model Evaluation

Assessing our model's accuracy,  
performance metrics, and limitations  
using real-world startup data.

# Our Data

## Features

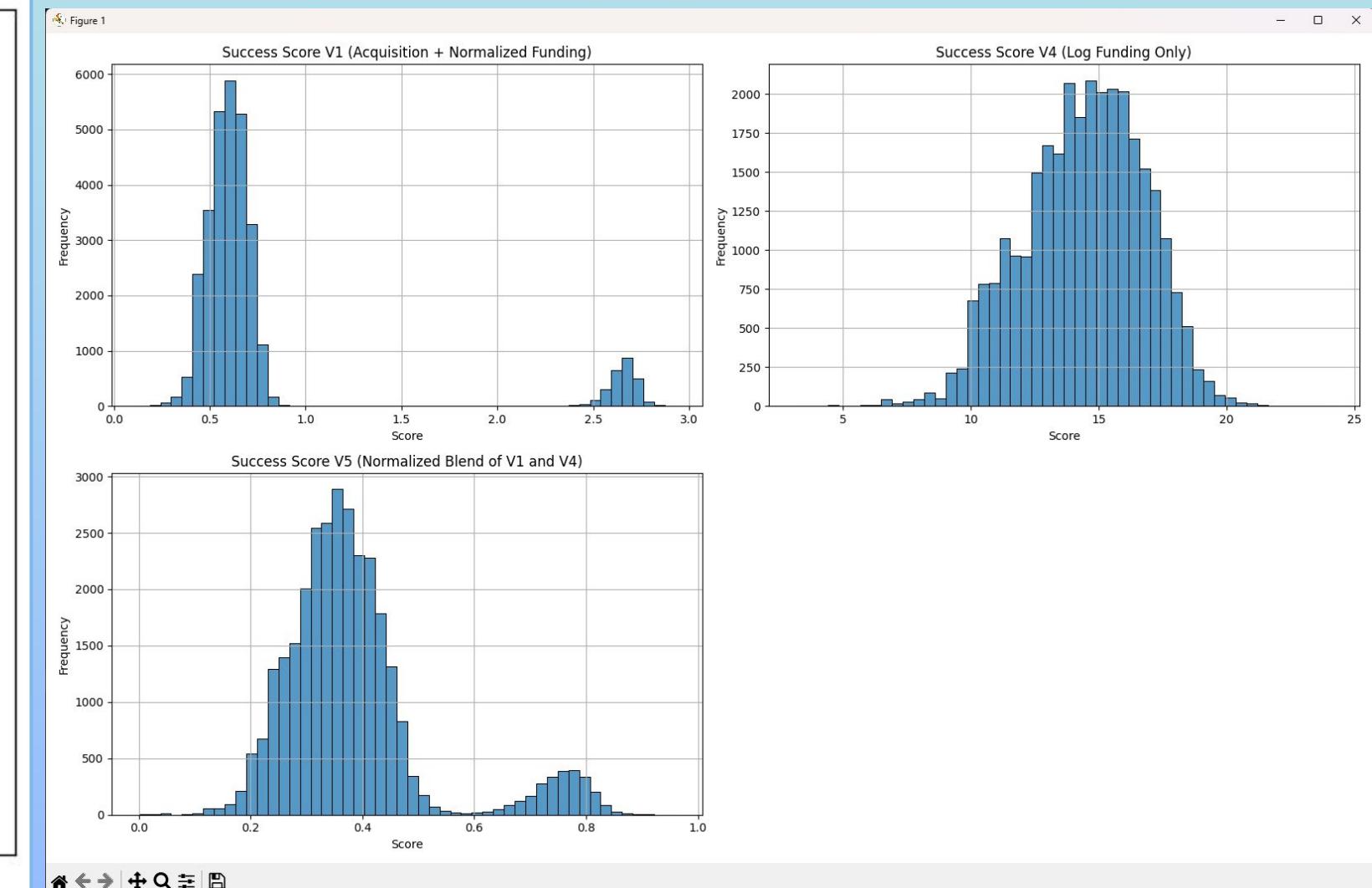
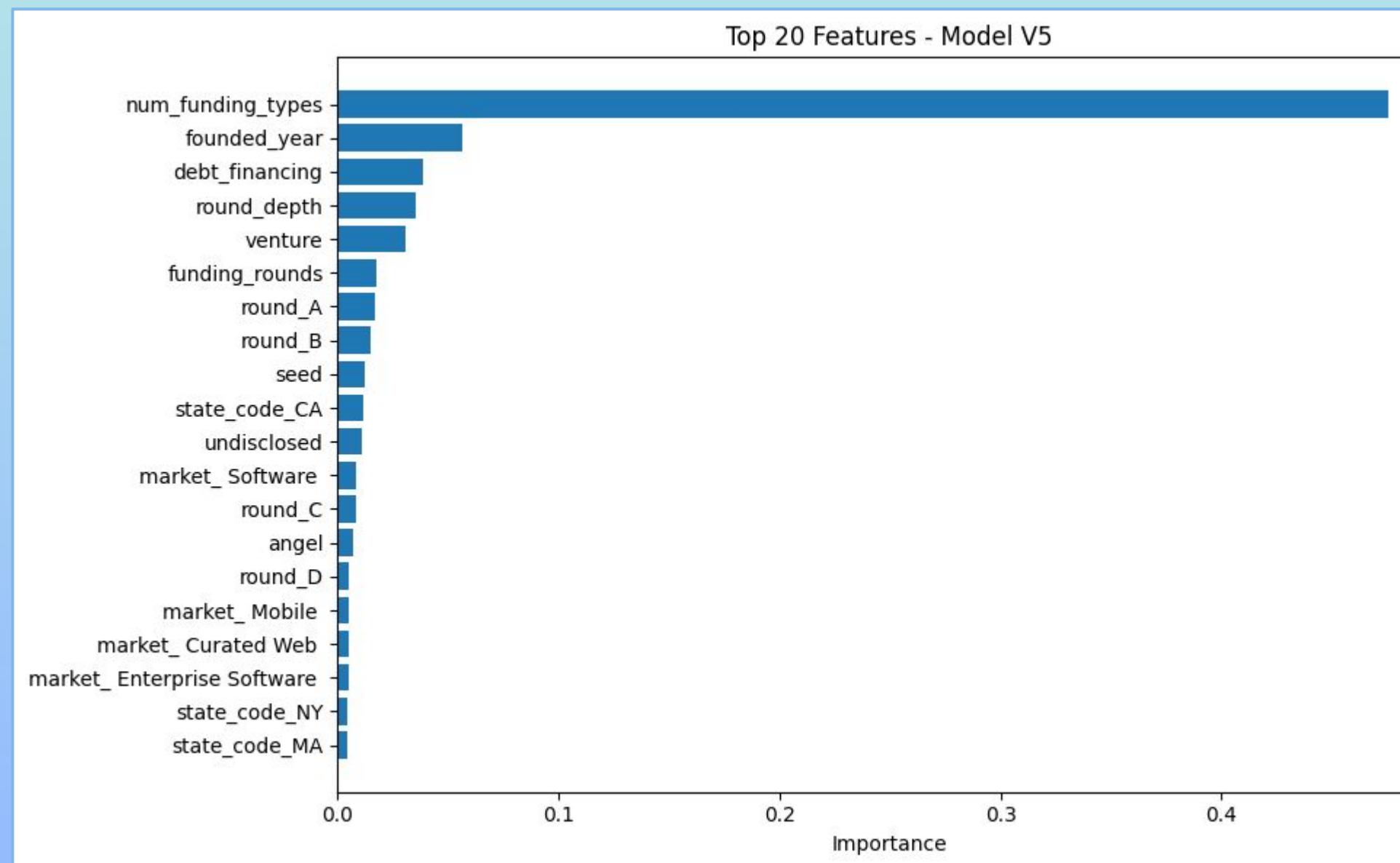
A total of 30 features in the original slide. We eliminated it to the 15 most highly correlated features for model training.

## Dataset

Our research utilized the [StartUp Investments \(crunchbase\)](#) dataset, which contains comprehensive data on startups from globally.

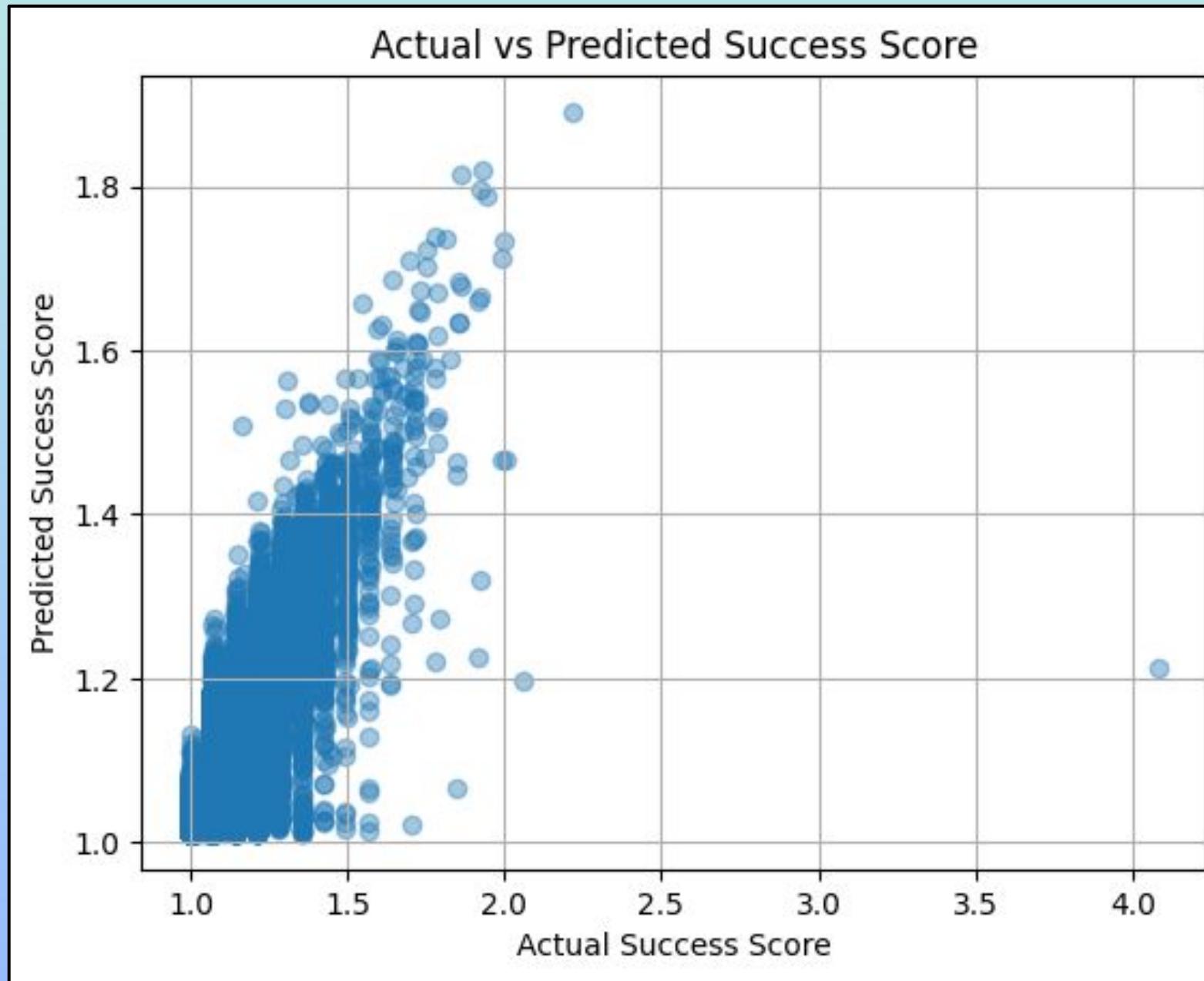
## Label

Success Score: incorporates multiple variables allowing for a detailed analysis of factors contributing to success.



# Model Metrics

We employed a supervised learning approach using the `RandomForestRegressor`. This method allowed us to analyze the relationship between various features and their predictive power regarding startup success.



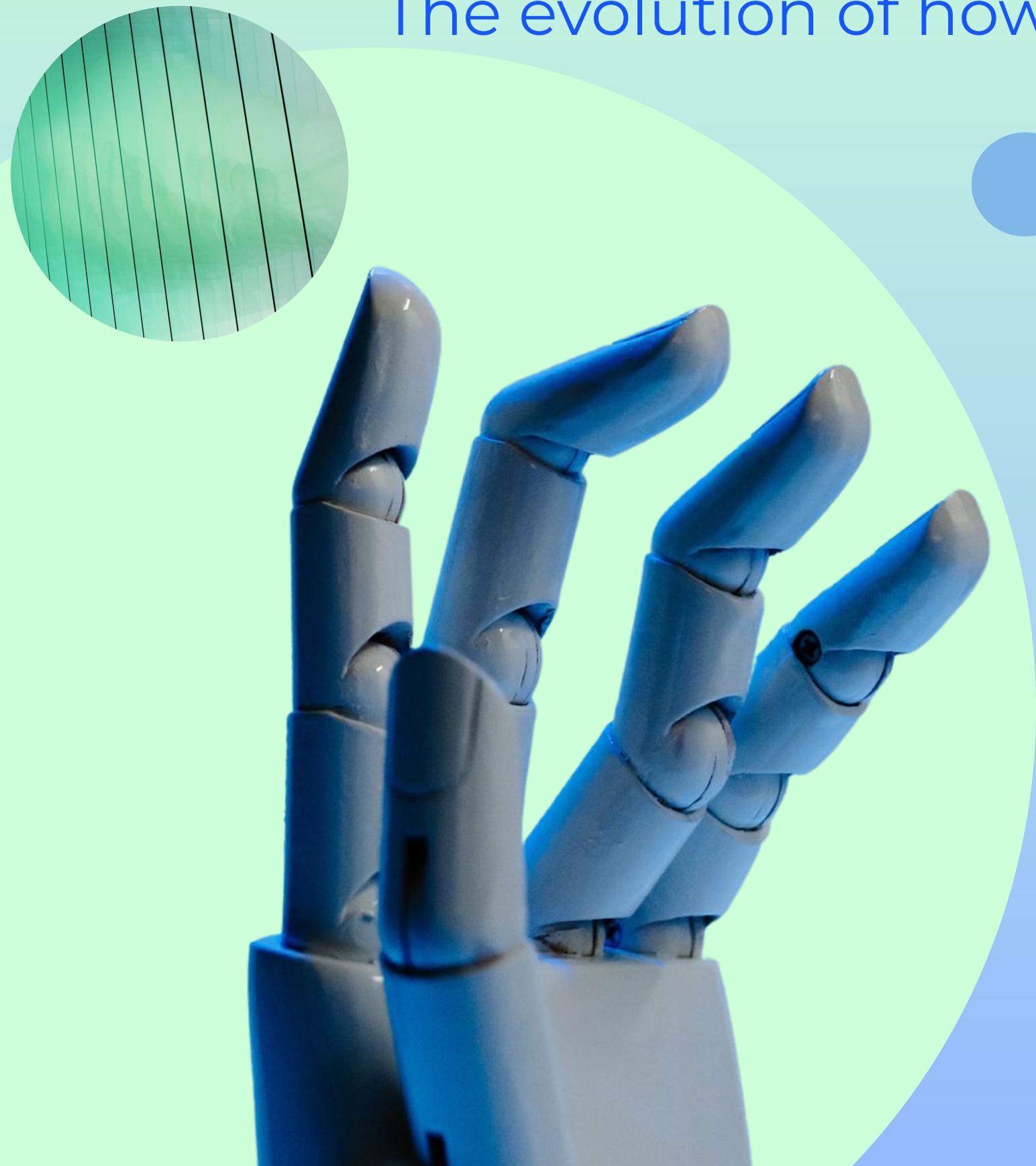
## Performance:

Our Random Forest Regression model achieved a **R<sup>2</sup> score of 0.4561**, indicating it explains about 46% of the variance in startup success scores.

The **Root Mean Squared Error (RMSE)** is **0.1017**, and the **Mean Absolute Error (MAE)** is **0.0490**, reflecting a good predictive accuracy with minimal deviation from true values.

# Our Success label

The evolution of how we defined & predicted “Success”



## Initial Dataset label

The dataset included a "**success score**" but it was noisy, poorly defined, and had little to no correlation.

## Creating our own score

We tried building a **custom success metric** using **all** of our features. The metrics were amazing but the model just learned the formula we encoded.

## Our Final Solution

We **redefined success** using just 2–3 core signals. We excluded these features from training to prevent data leakage

# SOURCES of BIAS

*How can the process of creating AI/ML solutions amplify or mitigate bias in the case of our group?*

1

## Time Frame Limitation

The dataset only includes startups founded latest in 2014, excluding later examples that could offer valuable insights.

2

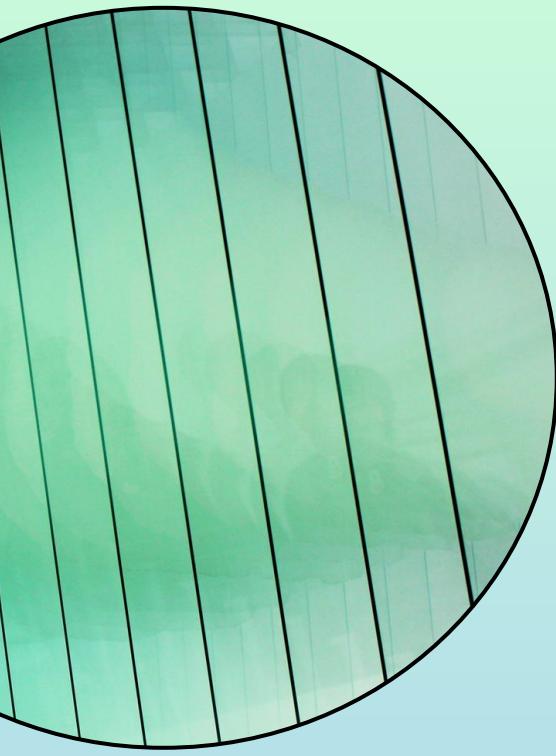
## Survivorship Bias

Failed startups, especially those that didn't receive funding, are underrepresented, which may skew the success patterns we observe.

3

## Nuanced Success Factors

Our definition of "success" is based on a few features and may vary across regions or industries.



# Next Steps & Ethical Considerations



## Expand the dataset

Integrate more diverse and recent startup data and include qualitative features like founder experience or team diversity.



## Explore alternative success factors

Revisit our "success" label → is it equitable across industries and contexts? How can we mitigate bias?



## UX/UI improvements

Tailor outputs to offer improved & constructive insights. Ensure accessibility and inclusivity in the platform design.

# Citations

- ➔ [Startup Statistics \(2025\) – Numbers By Country & Success Rate](#)
- ➔ [Global startup ecosystem value down, but AI-native start-ups on the rise – report](#)
- ➔ [Entrepreneurship, Startups, and Business Formation Are Booming Across the U.S.](#)



# Thank you!

