Utilizing Language Model Driven Founder Segmentation for Predicting Entrepreneurial Success

Anushka Suyal

March 24, 2024

Abstract

This research investigates the efficacy of utilising Large Language Models for feature engineering to forecast the success of early-stage enterprises. LLMs are employed to develop segmentation frameworks and utilise them to categorise founders based on their profile summaries. These categories are used as features for our model, and machine learning algorithms are subsequently applied to predict startup success. The analyses of success rates across segments and evaluation metrics for the machine learning models highlight the effectiveness of using the segmentation frameworks in forecasting entrepreneurial success. The results underscore the potential of utilising language models in venture capital decision-making processes.

1 Introduction

In the fast-paced world of venture capital (VC) investing, identifying high-growth startups with the potential for success is crucial for maximizing returns. With the limitations of human judgment becoming increasingly apparent, VC investors are turning towards machine learning (ML) models to enhance their decision-making processes. However, handling a vast array of features related to founder profiles presents challenges such as feature engineering complexities.

This research introduces an innovative approach that leverages Language Model Driven Founder Summarization and Segmentation to categorise founders into five distinct success levels (L1-L5) and ten personas (A-J). By utilizing datasets of successful and unsuccessful founders and employing a unique segmentation process, this study aims to improve the accuracy of founder success prediction in the VC investment domain.

Through a detailed analysis of founder profiles and success rates across different levels, this research investigates the effectiveness of using language models to assist VC investors in making informed investment decisions.

2 Methodology

The research was conducted in two phases. In the initial phase, the methodology focused on developing a level-based segmentation framework for founders to classify them based on their profile data. Subsequently, the study aimed to assess the success rates associated with each identified category.

In the second phase, a new segmentation approach based on personas was established, using the LLM to classify founders accordingly. Following this segmentation, a machine learning model was trained on a dataset containing both level and persona features assigned to each founder to predict their likelihood of success. Additional attributes of the founders were then incorporated into the dataset to refine the model, resulting in enhanced predictions of identifying successful companies.

2.1 Data Preprocessing

We utilised the same dataset as in previous studies[1], comprising 4029 profiles of successful founders who have founded companies over USD 500M and 5702 unsuccessful founders who failed to scale their companies comparably. From this initial dataset of 9731 founders, we selected a statistically representative sample of 600 individuals, with 300 samples from each category. Subsequently, we preprocessed the data to extract attributes concerning the founders, such as age, location, education, employment history, and various other features believed to influence the success of the founder.

2.2 Profile Summarisation

This process employs LLMs to take the preprocessed data for each founder as input, and generate a comprehensive summary based on it. A carefully crafted prompt is designed to prioritise the clear enumeration of all education and employment particulars provided by the founder. Furthermore, external sources are consulted to gather statistical information pertinent to the founder's profile, including referencing the QS World rankings for the universities attended by the founder.

2.3 Segmentation and Success Rate Analysis

2.3.1 Phase 1 – Level-based segmentation

Founder Segmentation

We deployed LLMs to generate a level-based segmentation designed to categorize founders based on various criteria, including the length of their professional experience, educational background, positions they have held, and the funding received for the businesses they have founded. The segmentation is as follows –

- Level 5 (L5): Entrepreneur whose company has 100M+ USD funding, or is going through IPO
- Level 4 (L4): Entrepreneur who has had a small to medium-size exit or has worked as an executive at a notable technology company
- Level 3 (L3): 10-15 years of technical and management experience (e.g. working at big tech and unicorn startups or having a PhD)
- Level 2 (L2): Entrepreneurs with a few years of experience or accelerator graduates
- Level 1 (L1): Entrepreneurs that are disconnected from tech circles or that have negligible experience but with large potential

Subsequently, we implemented prompt chaining, a technique in which the generated summaries were fed as input into a new prompt to segment the founder according to their profile summary and the previously established level segmentation. This process employed methodologies such as Chain-of-Thought (CoT) and few-shot prompting to enhance the segmentation accuracy and efficiency.

Success Rate Analysis

This entailed aggregating our findings for each level of the segmentation across both datasets. Subsequently, we computed the success rate for each level, reflecting the percentage of founders classified into a specific level who achieved success. As demonstrated in this[2] paper, the level segmentation outlined above was observed to improve the accuracy of predicting success.

2.3.2 Phase 2 – Persona-based Segmentation

Founder Segmentation

We outlined details of a new persona-based system of classifying founders, and prompted the LLM to generate a corresponding segmentation –

- A - Successful Serial Entrepreneur who built a public company
- B - Successful Serial Entrepreneur who had a company with exit over 100M USD
- C - Fresh Graduate Entrepreneur who founded a company immediately after or within a year of finishing university
- D - First time entrepreneur who founded a company within 3-5 years of finishing university
- E - Founder who holds a PhD degree, or is/was a researcher, or is/was a professor at a university

- F - First Time Non-tech Entrepreneur who has 10+ years of experience in non-tech industries
- G - Tech Executive at tech company/companies for over 10 years
- H - Engineer/executive at a unicorn company
- I - Founder who has a background as an engineer or product manager at a tech firm
- J - any others that are not in the above categories

Following the segmentation process, founders were assigned to one of the ten personas based on their summary utilizing methodologies like Tree-of-Thought (ToT) and few-shot prompting.

Upon further analysis, it was noted that the segments within the persona-based classification system exhibited overlaps. To enhance prediction accuracy, the Language Model was leveraged to generate binary outputs for questions pertaining to segments that may potentially overlap in the persona-based segmentation. These questions are as follows -

1. Has the individual ever held a role at a Unicorn company?
2. Has the individual ever held a software engineer/product manager role?
3. Does the individual possess a degree in a science/engineering subject?
4. Have any of the universities attended by the individual been ranked in the top 20 QS World ranking?

Subsequently, the dataset was expanded by incorporating features derived from the aforementioned questions, along with the output generated by the Language Model for these specific features.

Success Rate Analysis

This process entailed compiling the outcomes for each persona within the segmentation across both datasets and determining the success rates associated with each persona.

2.4 Success Prediction using ML Models

This involved training the data derived from the above segmentations to forecast startup success and assess the influence of the level-based and persona-based segmentation frameworks on improving prediction accuracy. Diverse machine learning classifiers, such as Logistic Regression, Linear Regression, and Gradient Boosting, were employed for this purpose.

Subsequently, additional binary features were introduced to the dataset alongside the categorical features, Level and Persona, to enhance the success predictions of the model. The machine learning models were then retrained based on the extended dataset.

We conducted an analysis of the following metrics obtained from the models –

- Accuracy: The measure of the proportion of correctly predicted outcomes among all predictions made by the machine learning model regarding the success or failure of a startup.
- Precision: The ratio of correctly identified positive predictions to the total predicted positives, reflecting the model's ability to avoid false positives in assessing startup success.
- AUC/ROC Score: The evaluation metric that quantifies the model's capability to distinguish between successful and unsuccessful startup predictions, crucial for assessing the model's predictive power.
- F1 Score: The harmonic mean of precision and recall, providing a balanced assessment of the model's performance in correctly categorizing startup success within the VC domain.
- True Positive Rate (TP rate): The proportion of actual positive instances correctly identified as positive by a classification model.

These metrics facilitated the comparison of model performance using different combinations of features, allowing us to analyse the effectiveness of the segmentations utilised.

3 Results

3.1 Profile Summarisation

When the preprocessed data for three founder profiles was input into the LLM and it was deployed to generate a summary for each profile, the outputs were as follows -

```
Profile 1:

Max Levchin, a 47-year-old male, currently resides in San Francisco,
    California. He pursued a Bachelor's degree in Computer Science at the
    University of Illinois Urbana-Champaign from September 1993 to May 1997.
    Levchin founded Affirm, where he serves as the Founder \& CEO since July
    2012. His employment history also includes roles such as Founder \&
    President at HVF Labs since January 2012 and Member Board Of Directors at
    Mixpanel from October 2016. Overall, he has been involved in various
    positions over the years. Levchin possesses skills in economics, software
    development, investment, and more.

1) The universities studied at, along with their global rankings according to
    the latest QS World University Rankings:

University of Illinois Urbana-Champaign, \#60

2) The jobs worked before founding companies:

VP of Engineering at Google, Google, American multinational technology company

Founder \& CEO at Slide, Slide, Software application developer

3) ALL the companies founded by them, status, net worth in USD:

Confinity, Acquired, \$1.5B

NetMeridian Software, Closed, \$0.5M

HVF Labs, Active, \$20.5M

Affirm, Active, \$2.0B



Profile 2:

Richard Malley, a 60-year-old male, currently resides in Boston,
    Massachusetts, United States. His education includes a PhD from Tufts
    University School of Medicine (1986-1990) and a Bachelor's in
    Psychobiology from Yale University (1982-1986). In terms of employment,
    Malley is the Scientific Founder and Board Member at Affinivax since June
    2014 and previously a Professor at Boston Children's Hospital from
    January 1990. His total work experience spans 31 years with 2 jobs.
    Malley's skills encompass medical research, healthcare, chemistry,
    economics, infectious diseases, clinical research, pediatrics, hospitals,
    research, teaching, and science.

1) The universities studied at, along with their global rankings according to
    the latest QS World University Rankings:

Tufts University School of Medicine, Not Ranked

Yale University, 17
```

```
2) The jobs worked before founding companies:

Professor, Boston Children's Hospital, Organization

3) All the companies founded by them in the format - company, status, total
    funding in USD:

Affinivax, Active, \$58.5M



Profile 3:

Thejo Kote, a 41-year-old male, currently resides in San Francisco,
    California, USA. He pursued a Master's degree in a 6-year program at UC
    Berkeley School of Information from January 2009 to January 2011 and a
    Bachelor's degree in Electrical Engineering at Visvesvaraya Technological
    University from January 2000 to January 2004. Kote's employment history
    includes roles such as Founder \& CEO at Airbase since November 2016,
    Board Member at TiE Silicon Valley from January 2018 to December 2019,
    and Co-founder \& CEO at Automatic from October 2010 to August 2016,
    among other positions. Overall, he has had a diverse professional
    experience contributing to different organizations over the years. Kote
    excels in skills related to programming languages, software development,
    product management, entrepreneurship, and more.

1) The universities studied at, along with their global rankings according to
    the latest QS World University Rankings:

Visvesvaraya Technological University, N/A

University of California Berkeley, N/A

2) The jobs worked before founding companies:

Software Engineer, Infosys, Indian multinational technology company

Senior Engineer, Netcore Cloud, Organization based in Mumbai, Maharashtra,
    India

3) ALL the companies founded/co-founded by them, status, net worth in USD:

Automatic, Acquired by SIRIUS XM HOLDINGS, \$1.0B

NextDrop, Active, Not Available
```

3.2 Segmentation and Success Rate Analysis

3.2.1 Phase 1 – Level-based segmentation

Founder Segmentation

The founder summaries were then analysed by the LLM, and each founder was segmented into one of the five levels L1-L5, based on our level segmentation framework -

```
Analyst Observations and Level Assignments:
Profile 1:
Analyst 1: Observation: Max Levchin has a strong educational background with
    a degree in Computer Science from a respectable university and a
    successful track record in founding companies, including Affirm, which
```

```
    has attained significant funding. He also has notable work experience at
    Google.
Level Assignment: L5 - Justification: Levchin meets the criteria for L5 as
    the founder and CEO of Affirm, a company with over 2.0 billion USD in
    funding, indicating substantial success and potential IPO readiness.

Analyst 2: Observation: Levchin has a successful entrepreneurial history,
    with the founding of companies like Confinity and Affirm, which have
    achieved substantial net worth. Additionally, his prior experience at
    Google in a leadership role adds to his qualifications.
Level Assignment: L5 - Justification: Levchin's track record of founding
    successful companies, including Affirm, which has amassed over 2.0
    billion USD in funding, aligns with the criteria for L5.

Analyst 3: Observation: Levchin's educational background, work experience at
    Google in a significant role, and successful track record in founding
    companies like Confinity and Affirm showcase a high level of expertise
    and achievement.
Level Assignment: L5 - Justification: Levchin's successful ventures,
    especially Affirm with over 2.0 billion USD in funding, demonstrate his
    capabilities as a top-tier entrepreneur on the verge of an IPO, meeting
    the requirements for L5.

In this analysis for Max Levchin, all three analysts independently assigned
    him to Level 5 (L5) due to his exceptional track record in founding
    successful companies like Confinity and Affirm, with the latter securing
    significant funding (2.0 billion USD) and indicating readiness for an
    IPO. Additionally, Levchin's educational background and work experience
    at Google further support his classification at Level 5.



Profile 2:
Analyst 1: Observation: Richard Malley possesses a Ph.D. from Tufts
    University School of Medicine and a Bachelor's from Yale University, a
    reputable institution. He has extensive experience in medical research
    and academia, including a position at Boston Children's Hospital. While
    he is the Scientific Founder of Affinivax, the total funding of 58.5
    million USD may not meet the criteria for L5. Level Assignment: L3 -
    Justification: Malley's educational background, extensive experience in
    medical research, and founding of Affinivax warrant a classification at
    Level 3 based on his technical and management expertise.
Analyst 2: Observation: Malley has a distinguished educational background
    with a Ph.D. and Bachelor's from respected universities. His experience
    as a Professor and Scientific Founder at Affinivax demonstrates expertise
    in medical research and healthcare. However, the total funding for
    Affinivax, while significant, might not reach the level required for L5.
    Level Assignment: L3 - Justification: Malley's strong educational
    foundation, experience in medical research, and founding of Affinivax
    with substantial funding place him at Level 3 due to the absence of
    evidence of a company with 100 million+ USD funding or IPO readiness.
Analyst 3: Observation: Malley's academic background, particularly with a
    Ph.D. and tenure as a Professor, reflects a high level of expertise in
    medical research. His role as the Scientific Founder of Affinivax, with
    58.5 million USD in total funding, showcases his entrepreneurial acumen.
    However, the funding amount may not reach the threshold for L5. Level
    Assignment: L3 - Justification: Based on Malley's academic achievements,
    research experience, and role in founding Affinivax, he aligns with Level
    3 given the substantial but not excessively high total funding of his
    founded company.
In this evaluation of Richard Malley's background, all three analysts
    independently assigned him to Level 3 (L3) based on his strong
```

Founder Segment	Total Count	Success Rate
L1	85	22%
L2	200	37%
L3	209	56%
L4	85	82%
L5	21	95%

Table 1: Success Rates for Level Segmentation.

```
    educational qualifications, extensive experience in medical research, and
    the founding of Affinivax. While his accomplishments and expertise merit
    a high level, the total funding of 58.5 million USD for Affinivax falls
    slightly below the threshold for Level 5, leading to a consensus
    classification at Level 3.



Profile 3:
Analyst 1: Observation: Thejo Kote holds both a Master's degree and a
    Bachelor's degree, indicating a strong educational background. He has
    founded companies like Automatic, which was acquired for 1.0 billion USD,
    showcasing significant entrepreneurial success. While the net worth of
    Automatic meets the L5 criteria, the absence of IPO or funding specifics
    may influence the level assignment.
Level Assignment: L4 - Justification: Kote's educational qualifications,
    successful founding of Automatic with a substantial acquisition value,
    and diverse work experience support a Level 4 classification based on his
    medium-size exit and executive roles
.
Analyst 2: Observation: Kote's educational achievements and founding of
    companies like Automatic with a notable acquisition value display a
    strong entrepreneurial track record. The lack of specific funding details
    for his companies may impact the level assessment.
Level Assignment: L4 - Justification: Given Kote's educational background,
    successful exit with Automatic, and diverse roles held, he fits the
    criteria for Level 4 as an entrepreneur with a medium-size exit and
    executive experience.

Analyst 3: Observation: Kote's educational attainments, particularly a
    Master's degree and founding of companies like Automatic, acquired for
    1.0 billion USD, demonstrate significant entrepreneurial prowess.
    However, the lack of information regarding IPO status or precise funding
    figures may influence the level determination.
Level Assignment: L4 - Justification: Considering Kote's educational
    qualifications, successful exit with Automatic, and overall professional
    experience, a Level 4 classification aligns with his background as an
    entrepreneur with a medium-size exit and executive roles.
In this assessment of Thejo Kote's profile, all three analysts independently
    assigned him to Level 4 (L4) based on his strong educational background,
    successful founding of Automatic (acquired for 1.0 billion USD), and
    diverse professional experience. While Kote's achievements meet the
    criteria for Level 5 in terms of company exit size, the lack of specific
    funding details for his companies led to a consensus classification at
    Level 4.
```

Success Rate Analysis

Table 1 shows the distribution for the level segments and their corresponding success rates.

3.2.2 Phase 2 – Persona-based Segmentation

Founder Segmentation The founder summaries were analysed by the LLM, and each founder was segmented into one of the ten personas (A-J), based on the persona segmentation. The following outputs were produced by the LLM -

```
Profile 1:

Here is the analysis for Max Levchin:

Founded Affirm in 2012 (15 years after finishing study in 1997): Does not
    meet criterion (C) or (D).
Co-founded multiple companies, including Affirm with a net worth of $5.0B,
    making him a serial entrepreneur with one company valued over $100M.
    Falls under class (B) - "Successful serial entrepreneur who had a company
    with exit over $100M".
His background in Computer Science and extensive experience in tech companies
    like Google point towards a tech persona, potentially class (G) - "Tech
    Executive at tech company/companies for over 10 years".
Reflecting upon these details, Max Levchin can be classified as:

Persona (G) "Tech Executive at tech company/companies for over 10 years."
This is supported by his educational and employment background in tech
    companies.

Final assignment: "G"



Profile 2:
Based on the founder summary provided for Richard Malley:

- Founded Affinivax in June 2014: Meets criterion (E) - "Founder holds a PhD
    degree or is/was a researcher, or is/was a professor at a University."
- Worked as a Professor at Boston Children's Hospital: Supports the
    classification under criterion (E).
- No information on the exit value of Affinivax available, so cannot
    categorize as (A) or (B).
- No indication of a public company or unicorn company affiliation, so does
    not qualify for (A) or (H).

Analysis:

Analyst 1 categorizes Richard Malley as Persona (E) - "Founder holds a PhD
    degree or is/was a researcher, or is/was a professor at a University."
    This is supported by his background in medical research and academia.

Analyst 2 agrees with this classification based on the clear evidence present
    in the founder summary.

Final assignment: "E"



Profile 3:

Based on the founder summary provided for Thejo Kote:

- Founded Automatic in October 2010, which was later acquired: Meets
    criterion (B) - "Successful serial entrepreneur who had a company with
    exit over $100M."
- Co-founded NextDrop in August 2009, which eventually closed: Does not meet
```

Founder Segment	Total Count	Success Rate
A	105	73%
B	62	61%
C	30	47%
D	119	25%
E	203	51%
F	11	18%
G	2	0%
H	9	67%
I	50	54%
J	9	22%

Table 2: Success Rates for Persona Segmentation.

```
    any specific criteria.
- Founded Airbase in November 2016: Does not provide enough information to
    qualify for an (A) level.
- Holds a Master's degree from UC Berkeley: Does not qualify for (E) as being
    a PhD holder or researcher.
- No clear tech executive role mentioned in his work history.

Reflecting on these details, Analyst 1 classifies Thejo Kote as:
Persona (B) "Successful serial entrepreneur who had a company with exit over
    $100M."
This is due to the successful exit of Automatic.

Analyst 2 argues Thejo Kote should be classified as:
Persona (J) "Any others that are not in the above categories."
This is because there is no concrete evidence for the other categories.

After deliberation, since Thejo Kote founded Automatic with a successful exit
    over $100M, both analysts agree that he falls under Persona (B) -
    "Successful serial entrepreneur who had a company with exit over $100M."

Final assignment: "B"
```

Success Rate Analysis

Table 2 presents the distribution of the persona segments and their corresponding success rates.

Based on the success rates for each persona segment, we can establish a correlation between the segment's description and the high success levels of the founders classified in them as follows:

- **A - Successful Serial Entrepreneur who built a public company (Success Rate: 73%):** Founders in this segment have proven track records of building successful companies and have likely developed valuable skills and networks over time. Their experience and expertise contribute to the high success rate observed.

- **B - Successful Serial Entrepreneur who had a company with exit over $100M (Success Rate: 61%):** Similar to segment A, founders in this segment have experience in building successful companies with significant exits, leading to a relatively high success rate. Their past successes and industry knowledge contribute to their success.

- **H - Engineer/executive at a unicorn company (Success Rate: 67%):** Founders with backgrounds at unicorn companies may have had exposure to successful business practices, innovative technologies, and high-growth environments. This background likely contributes to their higher success rate in entrepreneurship.

Model	**Accuracy**	**Precision**	**AUC-ROC**	**F1 Score**	**TPR**
Logistic Regression (L+P)	0.700	0.731	0.700	0.679	63%
Logistic Regression (with added features)	0.783	0.804	0.783	0.776	75%
Linear Regression (L+P)	0.783	0.758	0.783	0.794	81%
Linear Regression (with added features)	0.758	0.792	0.758	0.723	70%
Gradient Boosting (L+P)	0.667	0.685	0.667	0.649	62%
Gradient Boosting (with added features)	0.767	0.786	0.767	0.759	73%

Table 3: Metrics for Machine Learning Classifiers.

Test Data	**True Positive Rate**	**False Positive Rate**
2 successful, 28 unsuccessful	100%	20%
15 successful, 15 unsuccessful	100%	25%

Table 4: TPR and FPR for test data.

3.3 Success Prediction using ML Models

We trained Logistic Regression, Linear Regression, and Gradient Boosting machine learning classifiers using 'Level' and 'Persona' as categorical features to predict the 'Success' label. During the persona-based segmentation analysis using LLM responses, it was noted that some founders could be categorized into multiple segments, indicating non-mutually exclusive segments. To address this, four additional binary attributes were incorporated, reflecting whether the individual had worked at a unicorn company, held a software engineer/product manager role, obtained higher education from a top 20 ranked university, and possessed science/engineering degrees.

Subsequently, we computed the accuracy, precision, AUC/ROC score, F1 Score, and True Positive Rate for two models: one with 'Level' and 'Persona' as features (L+P in Table 3) and another model including L+P with the additional binary features (L+P with added features in Table 3).

4 Conclusion

4.1 Key Findings

Our study employed natural language processing to conduct founder segmentation based on profile summaries generated by large-language models using data extracted from the LinkedIn profiles of founders. The segmented data obtained from the classification was then utilized to train machine learning models aimed at predicting entrepreneurial success.

Significant improvements in accuracy were observed upon incorporating additional binary features derived from persona-segmentation overlaps. When utilising both level segmentation and persona segmentation to forecast success, Logistic Regression exhibited an initial accuracy of 70%. Upon the inclusion of new attributes, the accuracy increased by 8.3% to 78.3%, with precision rising by 7.3%. Gradient Boosting showcased a more considerable accuracy boost, soaring by 10%, from 66.7% to 76.7%, with precision also witnessing a notable surge of 10.1%, from 68.5% to 78.6%. Although Linear Regression experienced a slight 2.5% accuracy decline from 78.3% to 75.8%, the precision demonstrated a positive uplift of 3.4%.

These statistics indicated the effectiveness of the LLM-generated segmentation models and underscored the potential of leveraging language models in venture capital decision-making processes.

Upon deploying the trained Logistic Regression model to predict success on test data comprising 2 successful and 28 unsuccessful founders, the model accurately classified both successful founders, achieving a True Positive Rate of 100%. Subsequently, when the test data was adjusted to include 15 successful and 15 unsuccessful founders, the model correctly identified all 15 successful founders, achieving a True Positive Rate of 100%.

4.2 Limitations and Future Work

Limitations:

1. **Dataset Bias:** The success and failure criteria based on valuations of companies over USD 500M might introduce bias in the dataset, as success can be defined in various ways beyond monetary value. This could lead to skewed results.

2. **Overfitting:** The complexity of LLMs and the limited size of the dataset (600 founders), can lead to overfitting, where the model might learn to memorise patterns in the data rather than generalising well to new data.

3. **Model Stability:** The consistency of LLM responses may vary, leading to potential inconsistencies in the generated summaries and segmentations. This instability could impact the reliability of the model's output and decision-making process.

Future Work:

1. **Fine-tuning and Optimisation:** Conduct fine-tuning of the language model to enhance its performance in founder summarisation and segmentation tasks. Optimisation techniques such as hyperparameter tuning and regularization methods could improve the model's robustness and accuracy.

2. **Enhanced Persona Segmentation:** Refine the persona segmentation framework to introduce mutually exclusive segments that better capture the diversity of founder profiles. Incorporating additional criteria or refining existing segmentation parameters could lead to more precise segment assignments and higher classification accuracy.

References

[1] Sichao Xiong, Pantelis Papageorgiou, Fuat Varol Alican, and Yigit Ihlamur. Decision GPT: A Framework to Make Critical Decisions. 2023.

[2] Hao Li, Yigit Ihlamur. Large Language Model-Based Feature Engineering for Predicting Startup Success. 2024.