Startup Health Scoring Model Documentation

Sakshi Vedi
M.Sc. AIML, IIIT Lucknow

24 July 2025

---

1. Project Overview

This project is part of the ScaleDux AI Internship Task, focused on evaluating and ranking startups using a custom scoring system based on business and financial metrics. Each startup is scored out of 100 based on normalized features and assigned weights to reflect real-world investor logic.

2. Feature Weight Justification

The scoring formula was designed to reflect the most important indicators of a healthy startup. Weights were assigned to each feature as follows:

- **Market Size (25%)** – Indicates the potential of the product; a larger market gives room to grow.
- **Monthly Active Users (25%)** – Represents user traction and adoption.
- **Burn Rate (20%)** – A key risk factor; startups spending too fast may not sustain.
- **Team Experience (15%)** – Strong teams often pivot and execute better.
- **Funds Raised (15%)** – Helpful for scale, but not always a sign of performance.

3. Handling Burn Rate (Negatively Correlated Metric)

Burn rate is negatively correlated with performance — a higher burn means more cash outflow, which is typically a risk.

To account for this, we applied **Min-Max normalization** to all features, and then **inverted** the normalized burn rate:

$$\text{burn\_rate\_inverted} = 1 - \text{burn\_rate\_normalized}$$

This ensures that startups with lower burn get a higher contribution to their final score.

4. Observations and Insights

- Startups with high user traction and large market size consistently scored higher, even if funding was moderate.
- Some well-funded startups scored low due to high burn rate and poor user engagement.
- Efficient startups with limited funding but strong user growth and team experience outperformed overfunded peers.
- Correlation heatmaps revealed strong ties between market size and user base, but burn rate varied independently.

5. Deliverables

- Preprocessed and scored dataset in `cleaned_data/`
- Visualizations (bar chart, histogram, heatmap) in `outputs/`
- Colab notebook and modular Python scripts

Submitted by: Sakshi Vedi
M.Sc. AIML, IIIT Lucknow
Date: 24 July 2025