

**Project Proposal**

# **Early Stage Project Success Measurement**

**Presented By: Ahmed Hazmah Alfier Alshareef**

## **Content:**

- 1. Design**
- 2. Data & Algorithms**
- 3. Tools**
- 4. MVP**

# Design

---

## Introduction:

With the rapid changes in the global economy due to technology , tens if not hundreds of startups are launched every day. Those changes affected workers mindset because in one hand,some of them fear losing their jobs and on the other, some of them desire to impact the world which pushes towards starting new business. Thus, entrepreneurship is the mainstream in the educational systems as well as government and private sectors which can be observed clearly in Saudi Arabia in addition to other countries.

# Design

---

## Goal:

Therefore, the project goal mainly will help determine if an idea is worth pursuing or not (success or fail). Two main category of beneficiaries would use this project:

1. **Entrepreneurs: To assess the quality of their idea**
2. **Investors: To determine what startups to invest in.**

One way to measure the previous goal is by measuring user's interaction with the idea and how likely are the users to use it. The definition of success and failure is in the description of column "state"

# Data & Algorithms

---

## Data:

The dataset is from Kaggle and can be found here:

<https://www.kaggle.com/kemical/kickstarter-projects> , that contains the information of around 300,000 kickstarter projects. *kickstarter* is a crowdfunding platform.

# Data & Algorithms

---

**The dataset contains 13 columns as following:**

- ID: The project id
- name: The name of the project
- main category: The main category of the project (15 unique values)
- category: Sub category (158 unique values)
- currency: The currency of which the backers will be pay by (13 unique values)
- deadline: Last date of backing the project
- goal: The target money to launch the project in listed currency
- launched: The date of launching the project
- pledged: The amount of money that has been paid by backers to the project in listed currency
- state: Success if the project achieve at least the goal or failure otherwise
- backers: The number of people how backed the project
- country: The country of which the project is launched (21 unique values)
- usd pledged: Conversion in US dollars of the pledged column (done by kickstarter)

# Data & Algorithms

---

## Algorithms:

The goal as described yield a classification problem ;therefore, the model tries to predict the target column "state". Possible algorithms or architectures are Logistic Regression, Support Vector Machine, Random Forest, Gradient Boosting, XG Boosting, LSTM, and/or a transformer like "pre-trained Bert".

# Tools

---

**As the first glance at the dataset and the goal the following packages most likely will be used:**

1. **Data Processing:** Pandas, and Numpy
2. **Modelling:** SciKit-Learn, PyTorch, TensorFlow/Keras, and XGBoost. (Pre-trained models might be utilized)
3. **Visualization:** Matplotlib, and Seaborn (Plotly if needed)



# MVP

---

The minimum viable product (MVP) of this project is an exploratory data analysis (EDA) to check the factors contributing to the state of the project. The final product will contain a model to predict that state.