**Startup Unicorn Prediction Using Advanced**

**Machine Learning Algorithms**

**Abstract**

Artificial intelligence is an emerging topic and will soon be able to perform decisions better than humans. In more complex and creative contexts such as innovation, however, the question remains whether machines are superior to humans. Machines fail in two kinds of situations: processing and interpreting “soft” information (information that cannot be quantified) and making predictions in “unknowable risk” situations of extreme uncertainty. In such situations, the machine does not have representative information for a certain outcome. Thereby, humans are still the “gold standard” for assessing “soft” signals and make use intuition. To predict the success of startups, we, thus, combine the complementary capabilities of humans and machines in machine learning algorithms like Gradient Boosting, SVM, Random Forest and Decision Tree. All ML algorithms will be trained on past performance of STARTUP dataset and then this trained model can be used to predict success or failure of new STARTUP TEST DATA.

**Existing System**

One way towards understanding predictions in uncertain situations is to examine the mental processes that underlies the cognitive decision-making process. A theory that is particularly helpful in this context is the dual process theory of decision making. The underlying assumption of this theory is that people make use of two cognitive modes, one is characterized by intuition (system 1) and one by deliberate analytical predictions (system 2) (Tversky and Kahneman 1983; Kahneman 2011).

Predicting the success of early stage ventures is extremely complex and uncertain because frequently just vague ideas are prevalent, prototypes do not yet exist and thus the proof of concept is still pending. Moreover, such ideas might even not have a market yet, but offer great potential of growth in the future (Alvarez and Barney 2007). Consequently, the decision-making context is highly uncertain as neither possible outcomes nor the probability of such are known. This fact can be explained through two concepts: information asymmetry and unknowable risk (Alvarez and Barney 2007; Huang and Pearce 2015)

Information asymmetry describes situation, in which forecasters have incomplete information to decide (Spence 1974). When perfect information is absent, decision makers tend to search for various indicators that signal the likeliness of future outcomes (Morris 1987). In our context, such signals include both “hard” signals that can be easily quantified and categorized (e.g. industry, technology, team size) as well as “soft” signals (e.g. innovativeness, personality of entrepreneur). Humans then try to apply formal analysis to gather signals that support them in making deliberate, rule-based system 2 decisions (Kahneman 2011). On the other hand, unknowable risk defines situations in which a decision maker cannot gather information that signal a potential outcome or make decisions based on formal analysis because the simply not exist. This may be best compared to the error term of a statistical Bayesian model. Unknowable risk covers unexpected events that describe a deviation from status quo (Kaplan and March 1988). In our context, this means for instance identifying a unicorn startup that gains enormous return that only few would have expected. Formal analytics are not working in these contexts, as representative cases might be missing in previous experience. In such situations, where humans “don’t know what they don’t know”, decision making is mainly based on intuition (system 1) rather than formal analysis (Tversky and Kahneman 1983; Huang and Pearce 2015). Thus, predicting the success of early stage startups is a challenging task and the costs of misclassification are high as they might lead to disastrous funding decisions or missing valuable chances for return (Attenberg et al. 2015). Previous research in the context of early stage ventures provides strong evidence the best performance in terms of accuracy are provided by combining both types of predictions: analytical (system 2) and intuitive (system 1) (Huang and Pearce 2015).

**Disadvantages**

1.Less accuracy

**PROPOSED SYSTEM**

The objective of the project is to predict whether a startup which is currently operating turn into a success or a failure. The success of a company is defined as the event that gives the company's founders a large sum of money through the process of M&A (Merger and Acquisition) or an IPO (Initial Public Offering). A company would be considered as failed if it had to be shutdown.

To implement this project and predict success or failure of startup we have used advanced machine learning algorithms such as Gradient Boosting, SVM, Random Forest and Decision Tree. All ML algorithms will be trained on past performance of STARTUP dataset and then this trained model can be used to predict success or failure of new STARTUP TEST DATA.

**Advantages**

1.High accuracy

**SYSTEM REQUIREMENTS**

**HARDWARE REQUIREMENTS:**

# Processor - Pentium –IV

* Speed - 1.1 Ghz
* RAM - 256 MB(min)
* Hard Disk - 20 GB
* Key Board - Standard Windows Keyboard
* Mouse - Two or Three Button Mouse
* Monitor - SVGA

**SOFTWARE REQUIREMENTS:**

* Operating System - Windows7/8
* Programming Language - Python