# NITTE MEENAKSHI INSTITUTE OF TECHNOLOGY

(AN AUTONOMOUS INSTITUTION, AFFILIATED TO VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELGAUM, APPROVED BY AICTE & GOVT.OF KARNATAKA



# Start-Up Funding Analysis Report

*Submitted in partial fulfilment of the requirement for the award of Degree of*

*Bachelor of Engineering*

*in*

*Computer Science and Engineering*

*Submitted by:*

|  |  |
| --- | --- |
| Amith Kumar R | 1NT18CS009 |
| Monish K | 1NT18CS100 |
| Rithik G | 1NT18CS132 |

Under the Guidance of

DR. Vijaya Shetty S

Professor, Dept. of CS&E, NMIT

Department of Computer Science and Engineering

## (Accredited by NBA Tier-1)

2020-2021

## NITTE MEENAKSHI INSTITUTE OF TECHNOLOGY

(AN AUTONOMOUS INSTITUTION, AFFILIATED TO VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELGAUM

, APPROVED BY AICTE & GOVT.OF KARNATAKA)

Department of Computer Science and Engineering

## (Accredited by NBA Tier-1)



#### CERTIFICATE

This is to certify the Report on **“Indian-Startup Funding Analysis”** is an authentic work carried out by **Amith Kumar R(1NT18CS009), Monish K(1NT18CS100), Rithik G (1NT18CS132)** bonafide students of **Nitte Meenakshi Institute of Technology**, Bangalore in partial fulfilment for the award of the degree of ***Bachelor of Engineering*** in COMPUTER SCIENCE AND ENGINEERING of Visvesvaraya Technological University, Belagavi during the academic year ***2020-2021.*** It is certified that all corrections and suggestions indicated during the internal assessment has been incorporated in the report.

|  |  |
| --- | --- |
| **Internal Guide** | **Signature of the HOD** |
| DR. Vijaya Shetty  Professor, Dept. CSE, NMIT Bangalore | Dr.Thippeswamy M. N. Professor, Head, Dept. CSE, NMIT Bangalore |

We hereby declare that

## DECLARATION

1. The project work is our original work
2. This Project work has not been submitted for the award of any degree or examination at any other university/College/Institute.
3. This Project Work does not contain other persons’ data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This Project Work does not contain other persons’ writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
   1. their words have been re-written but the general information attributed to them has been referenced;
   2. where their exact words have been used, their writing has been placed inside quotation marks, and referenced.
5. This Project Work does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

|  |  |  |
| --- | --- | --- |
| **NAME** | **USN** | **Signature** |
| AMITH KUMAR R | 1NT18CS009 |  |
| MONISH K | 1NT18CS100 |  |
| RITHIK G | 1NT18CS132 |  |

### Date: 01/01/2021

**ACKNOWLEDGEMENT**

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of the people who made it possible, whose constant guidance and encouragement crowned our effort with success. I express my sincere gratitude to our Principal **Dr. H. C. Nagaraj**, Nitte Meenakshi Institute of Technology for providing facilities.

We wish to thank our HoD**, Dr.Thippeswamy M.N.** for the excellent environment created to further educational growth in our college. We also thank him for the invaluable guidance provided which has helped in the creation of a better project.

I hereby like to thank our ***DR.Vijaya Shetty, Professor***, Department of Computer Science & Engineering on her periodic inspection, time to time evaluation of the project and help to bring the project to the present form.

Thanks to our Departmental Project coordinators. We also thank all our friends, teaching and non- teaching staff at NMIT, Bangalore, for all the direct and indirect help provided in the completion of the project.

Signature

|  |  |  |
| --- | --- | --- |
| **NAME** | **USN** | **Signature** |
| AMITH KUMAR R | 1NT18CS009 |  |
| MONISH K | 1NT18CS100 |  |
| RITHIK G | 1NT18CS132 |  |

### Date: 1/1/2021

## ABSTRACT

The objectives of a startup are to be one’s own boss and to create employment to others which warrants lot of endurance and sacrifice.A startup is usually a company, such as a small business or an organization deliberately designed to swiftly develop a commercially viable business model.

More often than not, it is synonymous with high-tech projects, development and production, distribution of new products, processes or services. It usually comprises the design and execution of the innovative and out of the box processes of the development, validation and research for target markets.Large population with high percentage of middle income group, people with technical background, IT domination, high interest and mobile penetration are some of the drivers that have thrown up opportunities for spreading startup revolution in India.

In this project we analyze the start-up funding data across regions and sectors, and study the impact of venture capitalists for the creation of a startup.We make use of various graphical functions to show time series relationships in the dataset,build a KNN model to find the possible investors who can invest in the company in a city.The model will consider startups which belong to domains such as ecommerce, finance, healthcare, food, education, technology and logistics. This project work will help the potential enterprise to understand the current market economy which will allow them to predict the funding that they can expect from investors and also will help them to analyze the market and their chances of survival.

## TABLE OF CONTENTS

ACKNOWLEDGEMENT i

[ABSTRACT ii](#_TOC_250021)

[TABLE OF CONTENTS iii](#_TOC_250020)

[CHAPTER 1: INTRODUCTION 1](#_TOC_250018)

[CHAPTER 2: LITERATURE SURVEY 2-](#_TOC_250014)3

CHAPTER 3: SYSTEM REQUIREMENTS AND SPECIFICATIONS 4-4

* 1. HARDWARE REQUIREMENTS 4
  2. SOFTWARE REQUIREMENTS 4

[CHAPTER 4: IMPLEMENTATION](#_TOC_250010) 5-6

CHAPTER 5: RESULT 7

CHAPTER 6: CONCLUSION 8

BIBLIOGRAPHY 9

PLAGIARISM REPORT 10-11

## CHAPTER 1 INTRODUCTION

Indian Start-Up Funding Analysis analyzes the dataset which has funding information of the Indian startups from 2015 to 2020. It includes columns with the date funded, the city the startup is based out of, the names of the funders, and the amount invested (in USD).We build a KNN model to find the possible investors who can invest in the company in a city

This project will analyze the start-up funding data across regions and sectors, and study the impact of venture capitalists for the creation of a startup.We make use of various graphical functions to show time series relationships in the dataset,build a kmodel to find the possible investors who can invest in the company in a city.The model will consider startups which belong to domains such as ecommerce, finance, healthcare, food, education, technology and logistics. This project work will help the potential enterprise to understand the current market economy which will allow them to predict the funding that they can expect from investors and also will help them to analyze the market competition and their chances of survival.

Possible questions which could be answered are:

* How does the funding change with time?
* Do cities play a major role in funding?
* Which industries are favored by investors for funding?
* Who are the important investors in the Indian Ecosystem?
* How much funds does startups generally get in India?

## CHAPTER 2 LITERATURE SURVEY

A startup is a company that is in the first stage of its operations. These companies are often initially bankrolled by their entrepreneurial founders as they attempt to capitalize on developing a product or service for which they believe there is a demand. Due to limited revenue or high costs, most of these small-scale operations are not sustainable in the long term without additional funding from venture capitalists.

1. The recent trend towards the growth of start ups in India has indicated some fundamental patterns or commonalities between all the founders and starts ups. The recent trend which we are witnessing today can be track back to year 2010-11. This was the time when maximum number of start ups got established and still existing with their differentiated products/services and enjoying the first mover’s advantage. In a way one can say that the start up culture in India is just 7 years old and for that short period of time, India has done tremendously well of itself. One thing remains undone, which is to unlock or decode the pattern of founders/entrepreneurs.All the start ups India are largely based on IT/internet technology. Very few start ups are there which useother then IT for their operations.The most important element in Indian start ups is they are highly urban centric. Their operation and management remains in highly urbanized centres of country.[2] There is much uncertainty involved with early stage startups. Venture capitalists are faced with literally million dollar questions as they seek to evaluate early stage startups, to determine the potential of an investment. Likewise, entrepreneurs are faced with equally valuable questions as they seek venture capitalists from whom to not only gain funding, but also support and mentorship. We hope that this work will provide some answers regarding the VC-startup ecosystem and reveal key trends and indicators of success by studying the formal and informal networks of both venture capitalists and entrepreneurs.

The start-ups in India face many challenges in their early days few such are:-A large market

opportunity but Indians still do not have the discretionary income needed to create unparalleled

products. India’s middle class of about 78 million only earn INR 250,000 per year according to

the National Institute for Applied Economic Research (The Economist 2019) and only 50 million

have Shopped online (Google, Bain & Company and Omidyar Network 2019).

Financing still remains a challenge for Indian startups. While deal sizes of VC funding are significantly smaller than in Silicon Valley, considering India is an emerging economy, domestic lending rates are very high at three times those of developed economies, which incentivizes foreign funds to lend at cheaper rates.

1. The model will consider startups which belong to domains such as ecommerce, finance, healthcare, food, education, technology and logistics. This project work will help the potential enterprise to understand the current market economy which will allow them to predict the funding that they can expect from investors and also will help them to analyze the market competition and their chances of survival

[6]The methodology of data analysis comprises of 4 steps which are:-Data acquisition,Data Preprocessing,analyze data,result interpretation

Before doing any data mining to support the on-site decision making system , it is a necessity to do a minimum data preprocessing such as Data Classification , Data Cleaning , Data Transformation , Data Reduction [4]. All the data generated is not valuable for every use case. A vast proportion of data is always Dirty. Dirty data can significantly affect the decision making process Just by using tradition data cleaning process such as , noisy data processing ;  missing data processing ; redundant data processing ; duplicate data processing, might  be helpful but the data that needs to be feed to the system should always be Existent , Valid , Consistent , Integral , Accurate and Relevant [5].

1. The data is analyzed using python libraries like pandas, numpy, dateutil etc. Using these and many other libraries in-depth analysis was possible. The results are interpreted visually for a better understanding of what the data says. This visualization was possible by using python libraries like matlplotlib seaborn and squarify. Various visualization graphs and diagrams are used like bar plots and square plots. Bar plots are chats that represent categorical data with rectangular bar. These results would help us in understand where our startup ecosystem stands and what major factors play role in the funding of a startup

## CHAPTER 3

**SYSTEM REQUIREMENTS AND SPECIFICATIONS**

* 1. **HARDWARE REQUIREMENTS** PROCESSOR: Intel Core 2 Duo 1.3GHz or Faster. RAM:512MB or More.

VIDEO CARD: At least 128 MB of Video Memory. SPACE REQUIRED: 400MB.

* 1. **SOFTWARE REQUIREMENTS** OS:Windows XP,7,8,8.1,10, Linux,Ubuntu,Mac OS. ADDITIONAL SOFTWARE REQUIRED:- PYTHON 3

## CHAPTER 4 IMPLEMENTATION

**LIBRARIES USED:-**

* Pandas:- pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series.
* Matplotlib:-Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK+.
* Seaborn:-Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.
* Plotly:-Plotly's Python graphing library makes interactive, publication-quality graphs. Examples of how to make line plots, scatter plots, area charts, bar charts, error bars, box plots, histograms, heatmaps, subplots, multiple-axes, polar charts, and bubble charts.

**DATA:-**

Data was collected from kaggle website.Dataset consists of data of various start-up companies.

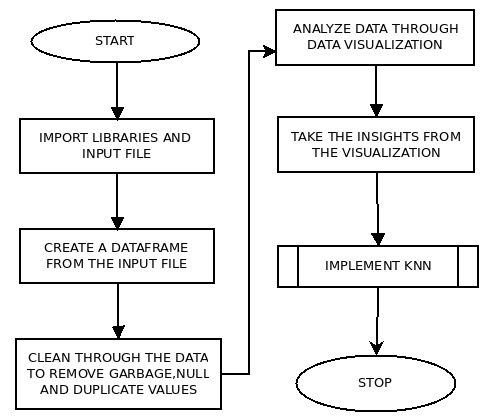
**Algorithm:-**

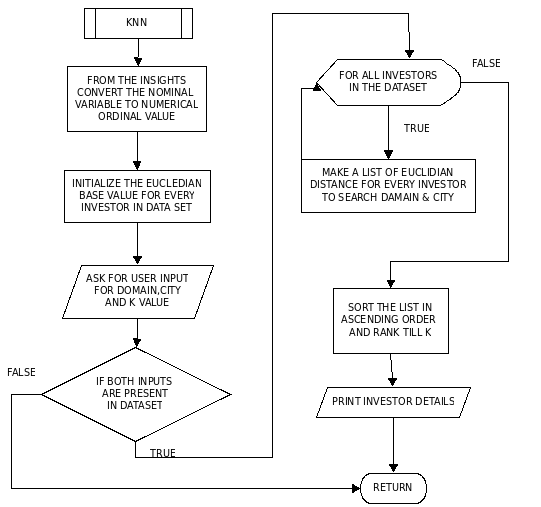
A modified version of knn algorithm is used, where the algorithm is modified to perform

Prediction rather than classification.

The KNN working can be explained on the basis of the below algorithm:

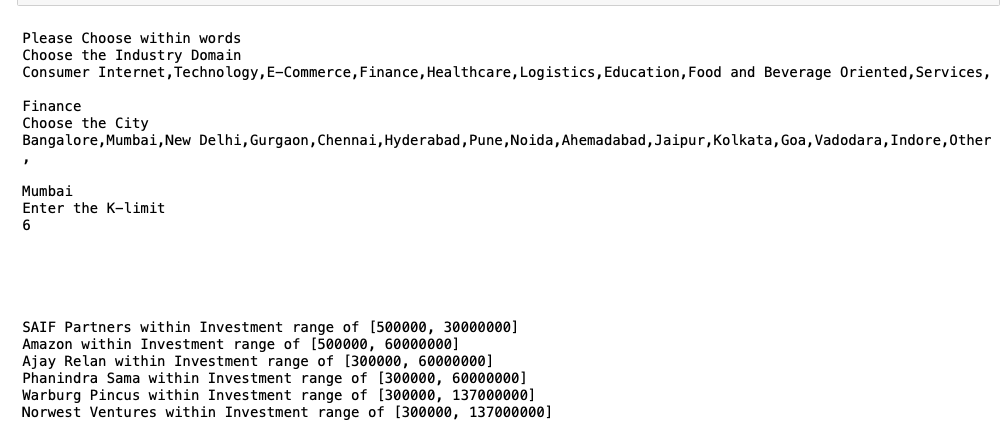
* **Step-1:** Select the number K of the neighbors
* **Step-2:** Calculate the Euclidean distance of **K number of neighbors**
* **Step-3:** Take the K nearest neighbors as per the calculated Euclidean distance.
* **Step-4:** Among these k neighbors, count the number of the data points in each category.
* **Step-5:** Assign the new data points to that category for which the number of the neighbor is maximum.
* **Step-6:** Our model is ready.





**CHAPTER 6**

**RESULT**



**CHAPTER 6 CONCLUSION**

A startup is usually a company, such as a small business or an organization deliberately designed to swiftly develop a commercially viable business model. In this project we have modified and implemented KNN algorithm which produce a list of investor who are more likely to invest. The probability of list of suitable investor generated is increased with the initial EDA performed through visualisation.we also analyze the start-up funding data across regions and sectors, and study the impact of venture capitalists for the creation of a startup.This project work will help the potential enterprise to understand the current market economy which will allow them to predict the funding that they can expect from investors and also will help them to analyze the market competition and their chances of survival.

**BIBLIOGRAPHY**

**Referenced books:-**

[7] Pang-Ning Tan, Vipin Kumar, Michael Steinbach: Introduction to Data Mining, Pearson, 2012

[8] Jiawei Han and Micheline Kamber: Data Mining - Concepts and Techniques, 3rd Edition, Morgan Kaufmann Publisher, 2014

**References**:-

1. Start Ups; Let’s Start Them Up - An Inside View in the Indian Start Up Scenario Dr. C Shekhar Upadhyay , Dr.Priyanka Rawal, Assistant Professor, JagranLakecity University, Bhopal
2. THE STARTUP ENVIRONMENT AND FUNDING ACTIVITY IN INDIA Dharish David, Sasidaran Gopalan, and Suma Ramachandran
3. An Analysis of Startup Ecosystem in Metropolitan City in India Neeraj K Pandey Academic Associate, Department of Management, Ahmedabad University, INDIA

[4] Gao Qiang, Wang Xinmin, Dong Chao, “Data Preparation for Prediction of Rerecirculating water chemistry Fault”,International Conference on Computer Application and System Modeling (ICCASM 2010) , 2010.

[5] Virender Kumar, Cherry Khosla, “Data Cleaning - A through Analysis and Survey on Unstructured Data”, 8th International Conference on Cloud Computing, Data Science and Engineering (Confluence), 2018.

[6] Data Analysis of startups investments and funding trends in India Piyush Anand Verma, Vikas Singha

**PLAGARISM REPORT**

