**Capstone-Project-2: Data Visualization and Inference Modeling-The Case of Nifty**

PROJECT REPORT

**Data Science with Python Programming**

**INDUSTRIAL PROJECT BASED LEARNING**



**Department of Computer Science and Engineering**

**Accredited by NBA**

**Geethanjali College of Engineering and Technology**

**(UGC Autonomous)**

(Affiliated to J.N.T.U.H, Approved by AICTE, New Delhi)

Cheeryal (V), Keesara (M), Medchal.Dist.-501 301.

**By Team -4**

**Team Details :**

Shivanoor Vignesh 21R11A0597

G.SandhyaRani 21R11A0569

R.Dhathri 21R11A0545

Hasini 21R11A0518

A.Pooja 21R11A0501

PSBN Sriya 21R11A0595

N Sai Charan 21R11A0588

MD.Akbar Ali 21R11A0586

# ABSTRACT

# This project delves into the analysis of stock market data with a focus on Nifty monthly and yearly returns, spanning a period of 20 years. The objective is to discern patterns, particularly emphasizing the year 2020, a significant period marked by the COVID-19 pandemic. Leveraging Python libraries such as Seaborn, Pandas, yfinance, and Matplotlib, the project entails importing the dataset, creating heat maps of monthly returns, and generating histograms to visualize the distribution of monthly and yearly returns across different standard deviation buckets. Additionally, an exploratory data analysis is conducted to derive insightful observations and inferences. Moreover, the project extends to building a web application using the Flask micro web framework to provide an interactive interface for users. This endeavor aims to enhance understanding of stock market dynamics and aid in making informed investment decisions.

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **Chapter No** | **Chapter Name** | **Page No** |
|  | **ABSTRACT** | 2 |
| **1** | **INTRODUCTION** | 4 |
| **2** | **LITERATURE SURVEY** | 5 |
| **3** | **PROBLEM STATEMENT** | 6 |
| **4** | **OBJECTIVES** | 7 |
| **5** | **METHODOLOGY** | 8 |
|  | **5.1 Data Source** | 8 |
|  | **5.2 Exploratory Data Analysis** | 10 |
|  | 5.2.1 Slicing and Dicing | 12 |
|  | 5.2.2 Checking for Data Consistency | 12 |
|  | 5.2.3 Data Visualization | 15 |
| **6** | **ALGORITHMS** | 26 |
| **7** | **IMPLEMENTATION** | 26 |
|  | 7.1 KNearNeighbour | 26 |
|  | 7.2 RandomForest Classifier | 27 |
|  | 7.3 Decision Tree Classifier | 28 |
| **8** | **Final Implementation** | 30 |
| **9** | **Final Observation** | 33 |
|  | **Conclusion** | 33 |

# INTRODUCTION

# The NIFTY 50 is a benchmark Indian stock market index that represents the weighted average of 50 of the largest Indian companies listed on the National Stock Exchange.

# The NIFTY 50 index has shaped up to be the largest single financial product in India, with an ecosystem consisting of exchange-traded funds (onshore and offshore), and futures and options at NSE and SGX. NIFTY 50 is the world's most actively traded contract. WFE, IOM and FIA surveys endorse NSE's leadership position. Between 2008 and 2012, the NIFTY 50 index's share of NSE market fell from 65% to 29%[10] due to the rise of sectoral indices like NIFTY Bank, NIFTY IT, NIFTY Pharma, and NIFTY Next 50.he Nifty experienced 15 crashes during the period 2000 to 2008 with a number of them having occurred in the months of January, May and June 2008. According to SEBI, approximately 89% of individual stock traders in the equity Futures & Options (F&O) segment incurred losses during the financial year

# 2020: COVID-19 Pandemic: In 2020, the world was hit by the COVID-19 pandemic, which caused widespread economic disruption and uncertainty. The pandemic led to lockdowns, travel restrictions, and disruptions in supply chains, which negatively impacted various sectors of the economy, including banking and finance.

# Economic Uncertainty: The uncertainty surrounding the duration and severity of the pandemic led to a sharp decline in economic activity. Banks faced challenges such as increased loan defaults, reduced consumer spending, and lower demand for credit.

# Market Volatility: The pandemic-induced market volatility caused panic selling and a flight to safety among investors. Banks, being sensitive to economic conditions and market sentiment, experienced significant declines in stock prices.

# 2008: Global Financial Crisis: The financial crisis of 2008 was triggered by the collapse of the subprime mortgage market in the United States, leading to a broader banking and financial crisis worldwide.

# Subprime Mortgage Crisis: Banks and financial institutions had invested heavily in mortgage-backed securities tied to subprime mortgages, which defaulted en masse as the housing bubble burst. This led to massive losses for banks and a liquidity crunch in financial markets.

# Credit Crunch: The collapse of major financial institutions and the freezing of credit markets led to a credit crunch, making it difficult for banks to borrow funds or extend credit to businesses and consumers. This further exacerbated the financial turmoil.

# Regulatory Failures: Regulatory failures and lax lending standards allowed the subprime mortgage bubble to inflate unchecked, leading to systemic risks that ultimately culminated in the financial crisis. The crisis prompted regulatory reforms aimed at strengthening oversight of the banking and financial sectors to prevent future crises.

# PROBLEM STATEMENT

# The effective management of investment portfolios hinges on the ability to discern opportune moments for stock entry or exit, a task contingent upon astute analysis of stock market data. In this context, the project endeavors to undertake a comprehensive examination of stock market data, focusing specifically on the Nifty index. The dataset contains 20years of stock price related data of nifty monthly returns data

# The dataset provided encapsulates two decades of stock price data, encompassing monthly and yearly returns of Nifty. Despite its breadth, the project prioritizes the year 2020, aiming to unveil intricate patterns amidst the backdrop of the COVID-19 pandemic.

# Objectives

# Import the dataset into Python to facilitate further analysis and exploration.

# Construct a heat map to visualize Nifty's monthly returns over the designated timeframe, providing insights into temporal fluctuations and trends.

# Categorize monthly returns into distinct buckets based on standard deviations and generate a histogram to illustrate the distribution of returns. This segmentation allows for a nuanced understanding of return variability and the prevalence of outlier events.

# Create a histogram depicting the distribution of yearly returns, similar to the monthly analysis, by categorizing returns into standard deviation-based buckets. This macroscopic view of Nifty's performance on an annual basis aids in identifying trends and anomalies.

# Develop a WEBUI or application using the Flask micro web framework to extend the project beyond mere analysis. This interactive platform enables stakeholders to engage with the analyzed data, facilitating a deeper understanding of Nifty's performance and supporting informed investment decisions.

# METHODLOGY

# sa