A Real time/Societal Research Project Report

"INSTAGRAM REACH ANALYSIS"

Submitted in Partial Fulfillment of the Academic Requirement for the Award of Degree of

BACHELOR OF TECHNOLOGY

in

Computer Science and Engineering
(Artificial Intelligence & Machine Learning)

Submitted By:

A.VAISHNAVI	**(22R01A6666)**
A.SAHITHYA	**(22R01A6667)**
E.SRUTHI	**(22R01A6679)**
K.PRANEETH	**(22R01A6692)**

Under the esteemed guidance of
Mr.B.Anil Kumar
Asst.prof.Dept of CSE(AI&ML)

CMR INSTITUTE OF TECHNOLOGY

(UGCAUTONOMOUS)

Approved by AICTE, Permanent Affiliation to JNTUH, Accredited by NBA and NAAC with A+ Grade

Kandlakoya(V),MedchalDist-501 401

2023-24

CMR INSTITUTE OF TECHNOLOGY

(UGCAUTONOMOUS)

Approved by AICTE, Permanent Affiliation to JNTUH, Accredited by NBA and NAAC with A+ Grade

Kandlakoya (V),Medchal Dist-501 401

www.cmrithyderabad.edu.in

.

CERTIFICATE

This is to certify that an Industry oriented Mini Project entitled with **"INSTAGRAM REACH ANALYSIS"** is being submitted by:

A.VAISHNAVI (22R01A6666)

To JNTUH, Hyderabad, in partial fulfillment of the requirement for award of the degree of B.Tech in CSE (Artificial Intelligence &Machine Learning) and is a record of a bonafide work carried out under our guidance and supervision. The results in this project have been verified and are found to be satisfactory. The results embodied in this work have not been submitted to any other University forward of any other degree or diploma

Signature of Guide	**Signature of Coordinator**	**Signature of HOD**
Mr.B.Anil Kumar	Mr.K.V.Balamurali Krishna	Prof.P.Pavan Kumar
(Asst.professor)	(Asst.professor)	(Head of Department)
Dept.Of.CSE(AI&ML)	Dept.Of.CSE(AI&ML)	Dept.Of.CSE(AI&ML)

ACKNOWLEDGEMENT

I am extremely grateful to **Dr. M Janga Reddy,** Director, **Dr. G.Madhusudhan Reddy** Principal and **Prof.P.Pavan Kumar** , Head of Department, Dept of Computer Science and Engineering (AI &ML), CMR Institute of Technology for their inspiration and valuable guidance during entire duration.

I am extremely thankful to our guide **Mr.B.Anil Kumar** Dept of Computer Science and Engineering , CMR Institute of Technology for their constant guidance, encouragement and moral support throughout the project.

I will be failing in duty if I do not acknowledge with grateful thanks to the authors of the references and other literatures referred in this Project.

I express my thanks to all staff members and friends for all the help and coordination extended in bringing out this Project successfully in time.

Finally, I am very much thankful to my parents and relatives who guided directly or indirectly for every step towards success.

A.VAISHNAVI (22R01A6666)

ABSTRACT

This study looks at Instagram to understand what makes some posts reach more people than others. We examine different factors like the type of post, its quality, the timing of when it's shared, the use of hashtags, and how much people interact with it. This helps us learn what makes Instagram posts more visible to a larger audience.Next, we build a model to predict how far future posts will reach.We use methods like ARIMA, Prophet, and LSTM to create this forecasting model, then test it tosee which one is the most accurate.

With these insights, social media managers and content creators can improve their Instagram strategy. They can use what we've found to plan their posts better, choose the right times to share, and use the best hashtags, ultimately reaching more people and engaging with a larger audience Social media users are growing constantly every year. Spontaneously interest in influencer marketing is also increasing. In influencer marketing, reach is important to estimate how many users have seen posts. Hashtags are the one of the most powerful way to increase reach. Specifically micro or nano influencers, whose followers are less than 10,000 followers, are in the spotlight for influencer marketing.

Their follower numbers are low, but their reach per posts is high and marketing cost is low. In this study, we reveal the use of hashtags and the relationship between reach rate in nano influencer's posts. We collected 232 nano influencer (1,000 -10,000 followers) accounts' meta-data and 113,247 hashtags used in their contents. By using regression analysis, we investigated the relationship between the number of hashtags and reach rate. Then by Term Frequency-Inverse Document Frequency (TF-IDF), we investigated the unique hashtags used in the reach rate quarter group. In conclusion, there was no relationship between the number of hashtags and reach rate. However, there were results that accounts using informative hashtags can be more effective to get reach than self-presenting hashtags.

LIST OF FIGURES

LIST OF FIGURES

1. INTRODUCTION

1.1 ABOUT PROJECT

Instagram is a major social media platform with over a billion active users. For businesses, influencers, and content creators, it's a powerful tool to connect with audiences, promote products, and build a brand. However, achieving high visibility on Instagram isn't always straightforward. The "reach" of a post—how many people see it—depends on many factors, from the content's appeal to the timing of its posting. Understanding these factors and knowing how to optimize them is key to a successful Instagram strategy.

This study explores the elements that affect Instagram reach. We aim to identify patterns and trends that can help users understand what drives visibility and how they can improve their posts' performance. By examining different aspects such as post type, quality, timing, hashtags, and engagement, we provide insights into what works best for reaching a larger audience.

Furthermore, we introduce a forecasting model to predict future Instagram reach. By analyzing past data and applying advanced statistical and machine learning techniques, we can estimate how future posts might perform. This predictive approach offers valuable guidance for planning content and improving reach.

Our goal is to help Instagram users—whether they are marketers, small business owners, or individual content creators—make informed decisions that lead to greater success on the platform. By the end of this study, readers will have a clearer understanding of Instagram reach and practical tools to enhance their own social media strategy.
Instagram stands as one of today’s most widely embraced social media platforms. Among its diverse user base, individuals leverage Instagram for professional purposes like business promotion, portfolio development, blogging, and content creation. Given its extensive user reach across various niches, Instagram continually evolves to enhance its offerings for content creators and its user community, striving to provide a better experience for all. This article will guide you through the process of conducting an Instagram Reach Analysis using Python. This analysis is aimed at assisting content creators in comprehending how to adjust and adapt to Instagram’s ongoing transformations over the long term.

"Instagram Reach, your gateway to expanding your online presence and connecting with a wider audience! Instagram Reach is a metric that measures the number of unique users who view your posts, stories, and profile. It's a key indicator of how far your content is spreading and how effectively you're building your brand on the platform.

With Instagram Reach, you can:

- Track your account's growth and engagement
- Understand what content resonates with your audience
- Identify opportunities to increase your visibility
- Refine your strategy to maximize your online impact.

Instagram was designed to capitalise on the ever improving technology crossover between smartphones and photography. Year on year new mobile devices focus on their built-in camera functions as a key selling point, mostly due to the increasing photo and short video-led world of social sharing.

Instagram has played a central role in this century's popular culture, with popular users dubbed 'influencers'. Influencers have become a key demographic for Instagram, and the app has opened up its platform with shopping and other marketing tools to make these popular users more successful.

Instagram has received some criticism for its role in this growing industry, as photoshopped and fake images are used to sell products. However, Instagram has added filters and tags to ensure that photoshopped items and product placement are correctly identified.

1.2.LITERATURE SURVEY

I. Instagram Reach Analysis and Prediction Rakesh Kundu 1 , Saraswat Ghosh 1 , Shivam Shreyansh 1 , Yash Yadav 1 Prof. Bhasker Rao2 1 Student, Department of CSE, Dayananda Sagar Academy of Technology and Management, Bangalore,

India

An analysis of Instagram reach and prediction of future reach is an important topic in the field of social

Media platform.

II. RAVI GUSAIN1 , SAKSHAM PATHAK2 1, 2 Department of Information Technology, Maharaja Agrasen Institute of Technology, Delhi, India.

Instagram is photo and video sharing platform having a huge influence on the reach of a person or a brand worldwide.

III..Dr. Daryl D. Green1*, Dr. Richard Martinez1, Amalan Kadja2, Lauran Evenson2 Lisa MacManus2, Stephanie Dirlbeck2

Social media is a great way to foster a personal relationship between a brand and its followers.

1.3. EXISTING SYSTEM:

1. Content Quality:

   - Post high-quality photos and videos.

   - Keep a consistent look or theme across your posts.

2. Timing and Frequency:

   - Post at times when your audience is most active.

   - Keep a consistent posting schedule.

3. Engagement:

   - Respond to comments and messages.

   - Engage with followers' content to build community.

4. Hashtags and Captions:

   - Use popular and relevant hashtags.

   - Write engaging captions to encourage interaction.

5. Collaborations:

   - Partner with influencers or other brands to reach more people.

These approaches are widely used but may not always guarantee consistent reach due to Instagram's changing algorithms and competitive environment. This study aims to find a more data-driven approach to understand and predict reach on Instagram.

DISADVANTAGES OF EXISTING SYSTEM:

•Cyber Bullying
•Features are not available for everyone
•Copyright issues
•Limited Customizations
•Privacy Concern

1.4.PROPOSED SYSTEM:

Our proposed system helps users increase their Instagram reach by analyzing data and predicting trends. Here's how it works:

1.Data Analysis:- We collect and study Instagram data, like posts, engagement (likes, comments, shares), hashtags, and posting times, to find what influences reach.

2.Identify Key Factors: - We identify the most important factors that affect Instagram reach, like post type, timing, and user interactions.

3.Forecasting Model: - We build a model to predict future reach using advanced techniques. This model learns from past data to forecast how well future posts might perform.

4.Practical Recommendations: - The system provides suggestions to improve reach, including the best times to post, effective hashtags, and ways to boost engagement.

5.Adaptable and Updated:- The system adapts to changes in Instagram's algorithm, ensuring that users get up-to-date recommendations.

6.Source Analysis: This analysis aims to discern the primary avenues through which your content is discovered on Instagram, categorized as follows:

- From Home: Content viewed by your followers in their feeds.
- From Hashtags: Content discovered via hashtag usage.
- From Explore: Content featured on Instagram’s Explore page.
- From Other: Content discovered through unclassified sources.

ADVANTAGES OF PROPOSED SYSTEM:

- Better Customer Relationships
- Followers Growth :
- Track Your Progress
- Evaluate your target audience in more detail
- Identify Trending Topics
- Powerful Advertising Platform
- Perform at the top of your industry
- Understand your target audience
- Save time and resources
- Plan your campaigns and strategy
- Low cost
- Simple to use
- Basic data and tools
- Detailed information

- It helps to increase conversion rates by reaching their target audience and making them aware of their product.
- Through this we can track metrics like views, likes, comments, and click-through rates to assess the effectiveness of your Reels.
- Track key website metrics, like total page views and bounce rate, to understand if your website content is valuable and engaging.
- When you know who your target audience is, you can make better decisions about everything from product development to pricing.
- It helps to increase conversion rates by reaching their target audience and making them aware of their product.
- Experts suggest that you use popular hashtags if you want bursts of visibility and Hashtags are a powerful tool for discovering trending topics and joining relevant conversations on Instagram.

2 .REQUIREMENT SPECIFICATIONS

2.1REQUIREMENT ANALYSIS

2.1.1HARDWARE REQUIREMENTS

- **System** : Pentium i3 Processor.
- **Hard Disk** : 500 GB.
- **Monitor** : 15"" LED
- **Input Devices** : Keyboard, Mouse
- **Ram** : 4 GB

2.1.2SOFTWARE REQUIREMENTS

- **Operating system** : Windows 10.
- **Coding Language** : Python
- **Software used** : Jupyter notebo

2.1 SPECIFICATION PRINCIPLES

2.2.1.SOFTWARE DESCRIPTION:

❖ Python:

Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently where as other languages use punctuation, and it has fewer syntactical constructions than other languages.

Python is Interpreted − Python is processed at runtime by the interpreter. You do not need to compile your program before executing it. This is similar to PERL and PHP.

Python is Interactive − You can actually sit at a Python prompt and interact with the interpreter directly to write your programs.

Python is Object-Oriented − Python supports Object-Oriented style or technique of programming that encapsulates code within objects.

Python is a Beginner's Language − Python is a great language for the beginner-level programmers and supports the development of a wide range of applications from simple text processing to WWW browsers to games.

History of Python

Python was developed by Guido van Rossum in the late eighties and early nineties at the National Research Institute for Mathematics and Computer Science in the Netherlands.

Python is derived from many other languages, including ABC, Modula-3, C, C++, Algol-68, SmallTalk, and Unix shell and other scripting languages.

Python is copyrighted. Like Perl, Python source code is now available under the GNU General Public License (GPL).

Python is now maintained by a core development team at the institute, although Guido van Rossum still holds a vital role in directing its progress.

PYHTON FEATURES:

Python's features include −

- **Easy-to-learn** − Python has few keywords, simple structure, and a clearly defined syntax. This allows the student to pick up the language quickly.
- **Easy-to-read** − Python code is more clearly defined and visible to the eyes.
- **Easy-to-maintain** − Python's source code is fairly easy-to-maintain.
- **A broad standard library** − Python's bulk of the library is very portable and cross-platform compatible on UNIX, Windows, and Macintosh.
- **Interactive Mode** − Python has support for an interactive mode which allows interactive testing and debugging of snippets of code.
- **Portable** − Python can run on a wide variety of hardware platforms and has the same interface on all platforms..
- **Databases** − Python provides interfaces to all major commercial databases.
- **GUI Programming** − Python supports GUI applications that can be created and ported to many system calls, libraries and windows systems, such as Windows MFC, Macintosh, and the X Window system of Unix.
- **Scalable** − Python provides a better structure and support for large programs than shell scripting.

Apart from the above-mentioned features, Python has a big list of good features, few are listed below −

- It supports functional and structured programming methods as well as OOP.
- It can be used as a scripting language or can be compiled to bytecode for building large applications.
  It provides very high-level dynamic data types and supports dynamic type checking.
- It supports automatic garbage collection.

PYTHON INSTALL:

Many PCs and Macs will have python already installed.

To check if you have python installed on a Windows PC, search in the start bar for Python or run the following on the Command Line (cmd.exe):

C:\Users\Your Name>python --version

To check if you have python installed on a Linux or Mac, then on linux open the command line or on Mac open the Terminal and type:

python --version

If you find that you do not have python installed on your computer, then you can download it for free from the following website: **https://www.python.org/**

 Modules Which are being used for data exploration, pro processing and for using random forest algorithm are:

Python's influence extends far beyond its role as a mere programming language; it stands as a transformative force in various industries and spheres, underpinned by its remarkable simplicity and adaptability. Its pervasive appeal lies in its readability and accessibility, drawing both seasoned developers and novices from diverse fields into its fold.

JUPYTER NOTEBOOK:

- Jupyter Notebook is an incredibly powerful tool for interactively developing and presenting data science projects. It combines code, visualizations, narrative text, and other rich media into a single document, creating a cohesive and expressive workflow.

- At its core, a notebook is a document that blends code and its output seamlessly. It allows you to run code, display the results, and add explanations, formulas, and charts all in one place. This makes your work more transparent, understandable, and reproducible.

- Jupyter Notebooks have become an essential part of the data science workflow in companies and organizations worldwide. They enable data scientists to explore data, test hypotheses, and share insights efficiently.

- As an open-source project, Jupyter Notebooks are completely free. You can download the software directly from the Project Jupyter website or as part of the Anaconda data science toolkit.

***Users of the notebook online application can:**

- Automatic syntax highlighting and indentation are available when editing code in the browser.
- Activate the browser's code.
- Check out the computations' output in media formats like HTML, LaTex, PNG, PDF, etc.
- Create and use widgets in JavaScript.
- Contains mathematical formulas presented in Markdown cells

1.Kernels:

The independent processes launched by the notebook web application are known as kernels, and they are used to execute user code in the specified language and return results to the notebook web application.

The following languages are available for the Jupyter Notebook kernel:

1. Python
2. R
3. Julia
4. Ruby
5. Scala
6. node.js

2. Notebook documents

All content viewable in the notebook online application, including calculation inputs and outputs, text, mathematical equations, graphs, and photos, is represented in the notebook document.

JUPYTER INSTALL:

Getting Started with Jupyter Notebook

The easiest way to install jupyter notebook is through the terminal:

Step 1: Python's latest version for this method(https://www.python.org/downloads/).

Step 2 : Updating pip using cmd.

python -m pip install --upgrade pip

Screenshot-2023-09-13-205527

Upgrading pip

Step 3: Install the jupyter notebook using the command pip install jupyter notebook in the terminal.(refer to the image)

pip install jupyter notebook

Step 4: Use the command jupyter notebook in terminal to run the notebook.

jupyter notebook

After you type the command, this home page should open up in your default browser.

3.SYSTEM DESIGN

3.1.ARCHITECTURE DIAGRAM

The system architectural design is the design process for identifying the subsystems making up the system andframework for subsystem control and communication. The goal of the architectural design is to establish the overall structure of software system.

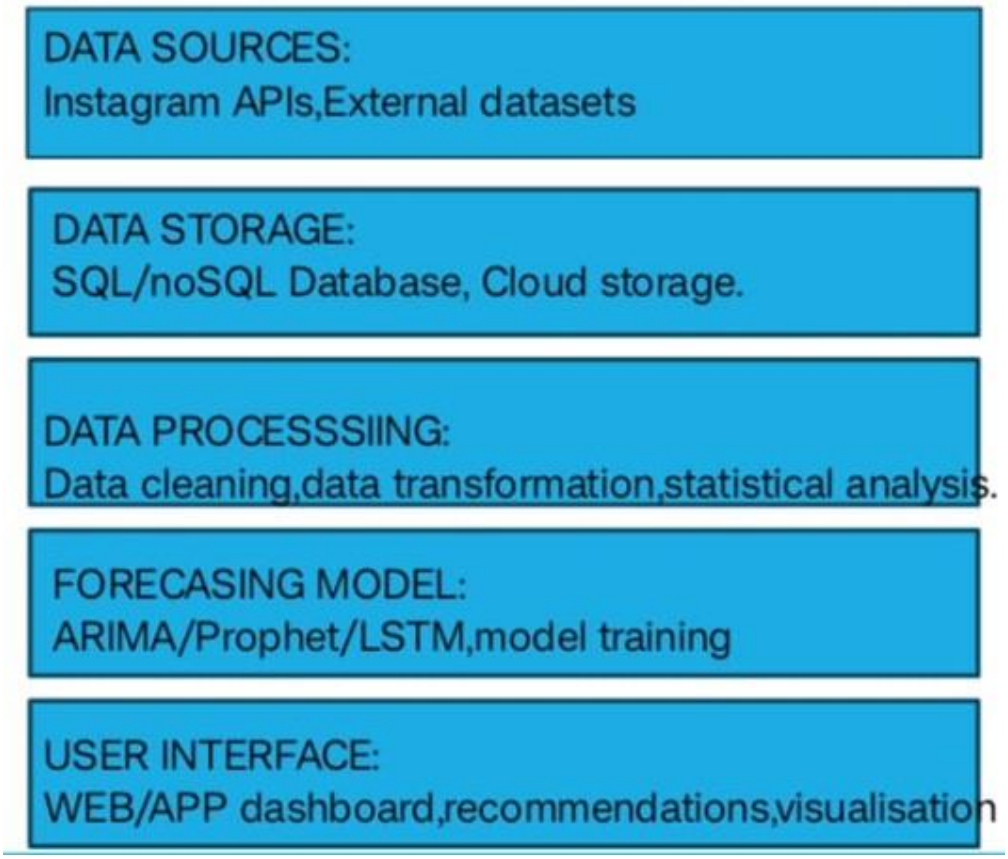


Fig.3.1.Architecture Diagram

1.1 UML DIAGRAMS

UML stands for Unified Modeling Language. UML is a standardized generalpurpose modeling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group.

The goal is for UML to become a common language for creating models of object oriented computer software. In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML.

The Unified Modeling Language is a standard language for specifying,

Visualization, Constructing and documenting the artifacts of software system, as well as for business modeling and other non-software systems.

The UML represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems.

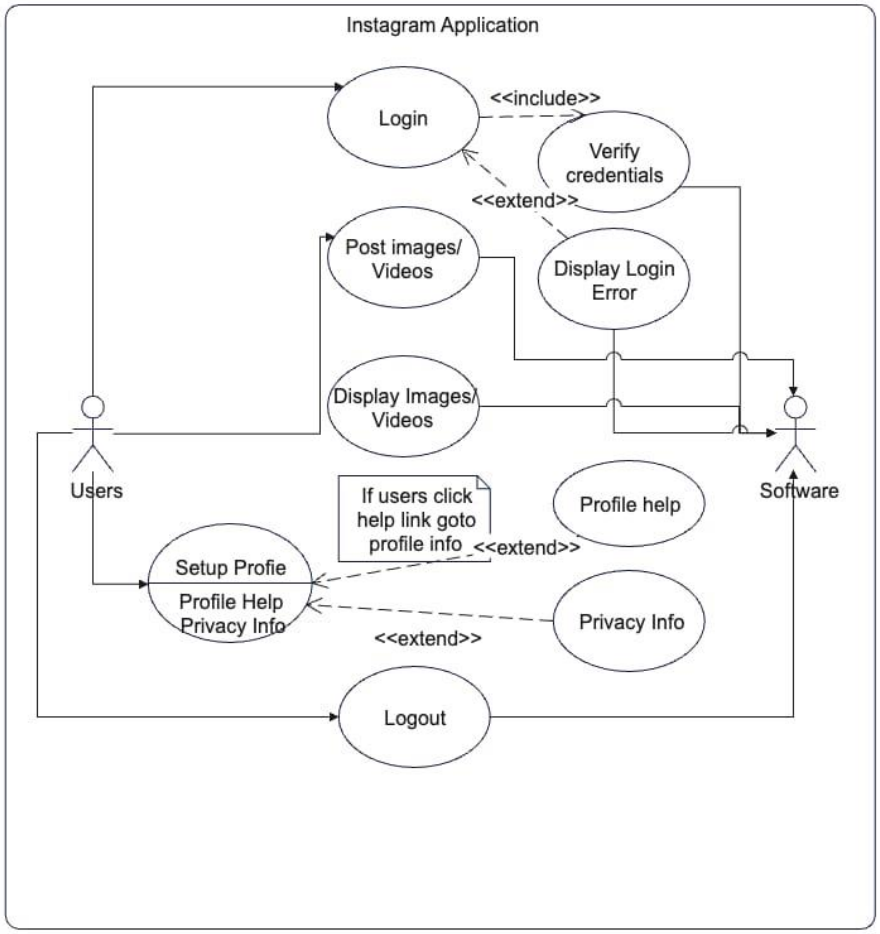
The UML is a very important part of developing objects oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

GOALS:

The Primary goals in the design of the UML are as follows:

1. Provide users a ready-to-use, expressive visual modeling Language so that they can develop and exchange meaningful models.
2. Provide extendibility and specialization mechanisms to extend the core concepts.
3. Be independent of particular programming languages and development process.
4. Provide a formal basis for understanding the modeling language.
5. Encourage the growth of OO tools market.
6. Support higher level development concepts such as collaborations, frameworks, patterns and components.

UML DIAGRAM:


Instagram Application
Login
<<include>>
Verify credentials
<<extend>>
Post images/ Videos
Display Login Error
Display Images/ Videos
Users
Software
If users click help link goto profile info
<<extend>>
Profile help
Setup Profie
Profile Help
Privacy Info
Privacy Info
<<extend>>
Logout

4.IMPLEMENTATION

4.1.PROJECT MODULES

4.1.1.MODULES:

- Data Collection
- Dataset
- Data Preparation
- Model Selection
- Analyze and Prediction
- Accuracy on test set
- Saving the Trained Model

4.1.2.MODULES DESCSRIPTION:

Data Collection:

This is the first real step towards the real development of a machine learning model, collecting data. This is a critical step that will cascade in how good the model will be, the more and better data that we get, the better our model will perform.
There are several techniques to collect the data, like web scraping, manual interventions and etc

Predictive Analysis for Big Mart Sales Using Machine Learning Algorithms

Data Preparation:

Wrangle data and prepare it for training. Clean that which may require it (remove duplicates, correct errors, deal with missing values, normalization, data type conversions, etc.

Model Selection:

We used decision tree regression machine learning algorithm , We got a accuracy of 95.7% on test set so we implemented this algorithm

4.1.ALGORITHM

4.2.2 RANDOM FOREST ALGORITHM

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model.

As the name suggests, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset."

Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.

The greater number of trees in the forest leads to higher accuracy and prevents the problem of over fitting.

The below diagram explains the working of the Random Forest algorithm:

Random Forest is a robust method for data analysis that excels in both classification and regression tasks. It operates by constructing multiple decision trees during training, offering powerful insights into feature importance and predictive accuracy.

In the context of data analysis, Random Forest proves effective for various purposes. For classification, it can discern patterns within categorical data, like identifying fraud transactions or predicting customer churn. In regression tasks, it can predict continuous variables, such as sales forecasting or housing prices.

Assumptions for Random Forest Since the random forest combines multiple trees to predict the class of the dataset, it is possible that some decision trees may predict the correct output, while others may not. But together, all the trees predict the correct output. Therefore, below are assumptions for a better Random forest classifier.

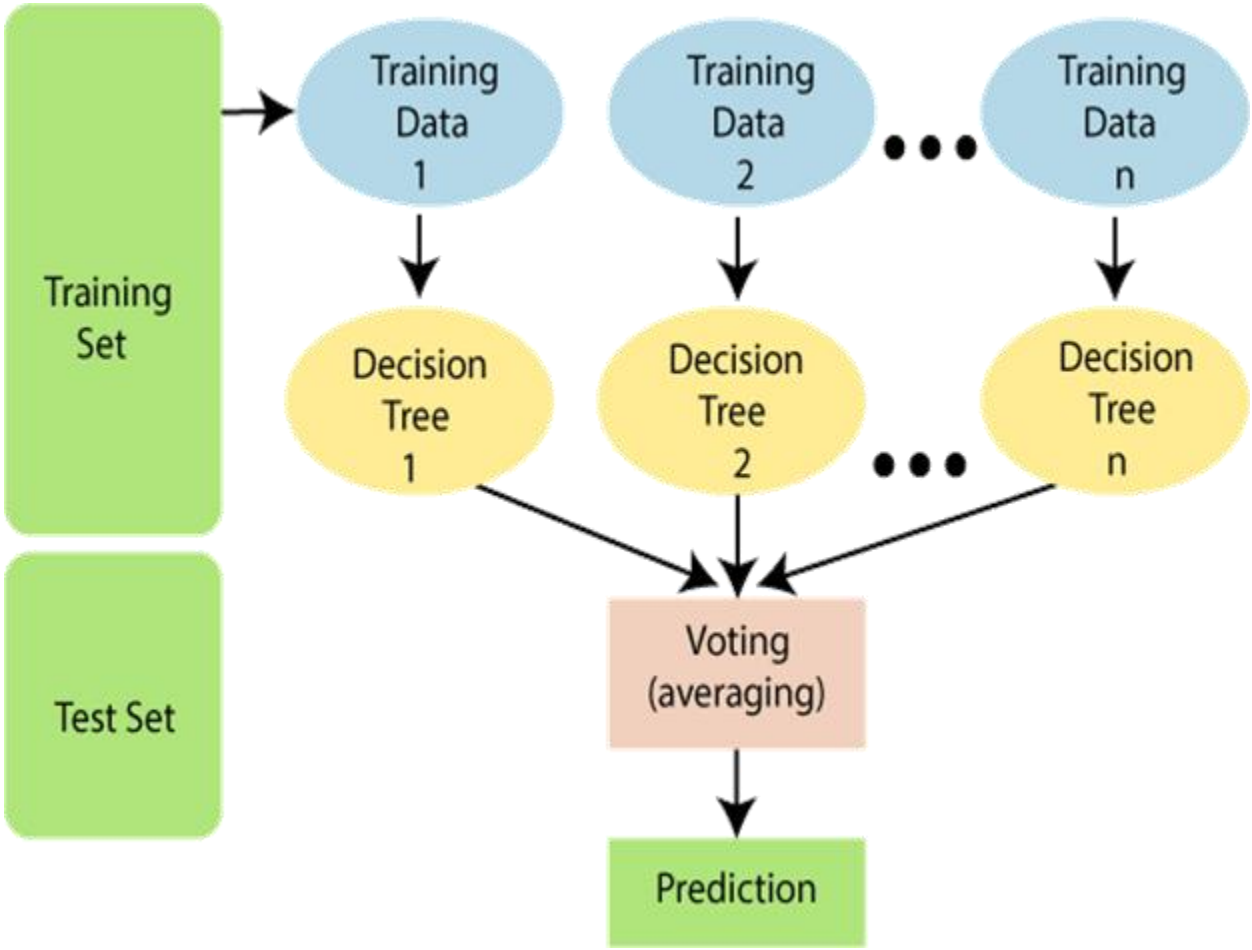


Fig.4.2.2.1 working of the Random Forest algorithm

Why use Random Forest?

Below are some points that explain why we should use the Random Forest algorithm:

<="" li="">

It takes less training time as compared to other algorithms.

- It predicts output with high accuracy, even for the large dataset it runs efficiently.
- It can also maintain accuracy when a large proportion of data is missing.

How does Random Forest algorithm work?

The Working process can be explained in the below steps and diagram:

Step-1: Select random K data points from the training set.

Step-2: Build the decision trees associated with the selected data points (Subsets).

Step-3: Choose the number N for decision trees that you want to build.

Step-4: Repeat Step 1 & 2.

Step-5: For new data points, find the predictions of each decision tree, and assign the new data points to the category that wins the majority votes.

Applications of Random Forest

There are mainly four sectors where Random forest mostly used:

1. **Banking:** Banking sector mostly uses this algorithm for the identification of loan risk.
2. **Medicine:** With the help of this algorithm, disease trends and risks of the disease can be identified.
3. **Land Use:** We can identify the areas of similar land use by this algorithm.
4. **Marketing:** Marketing trends can be identified using this algorithm.

Advantages of Random Forest

- Random Forest is capable of performing both Classification and Regression tasks.
- It is capable of handling large datasets with high dimensionality.
- It enhances the accuracy of the model and prevents the overfitting issue.
- **High Predictive Accuracy**: Outperforms single models, providing high accuracy in classification and regression tasks.
- **Feature Importance Insights**: Offers clarity on critical variables within the data for informed decision-making.
- **Robust to Noise**: Maintains reliability in complex datasets by effectively handling noisy or outlier-laden data.
- **Efficiency on Large Datasets**: Retains accuracy and efficiency even in extensive datasets, making it suitable for big data analysis.
- **User-Friendly**: Requires minimal parameter tuning and limited preprocessing, ensuring ease of use.

Disadvantages of Random Forest

- Although random forest can be used for both classification and regression tasks, it is not more suitable for Regression tasks.

4.3.SAMPLE CODE:

```
import pandas as pd
import plotly.graph_objs as go
import plotly.express as px
import plotly.io as pio
pio.templates.default = "plotly_white"

data = pd.read_csv("Instagram-Reach.csv", encoding = 'latin-1')
print(data.head())
```

Output:

	Date	Instagram reach
0	2022-04-01T00:00:00	7620
1	2022-04-02T00:00:00	12859
2	2022-04-03T00:00:00	16008
3	2022-04-04T00:00:00	24349
4	2022-04-05T00:00:00	20532

```
data['Date'] = pd.to_datetime(data['Date'])
print(data.head())
```

Output:

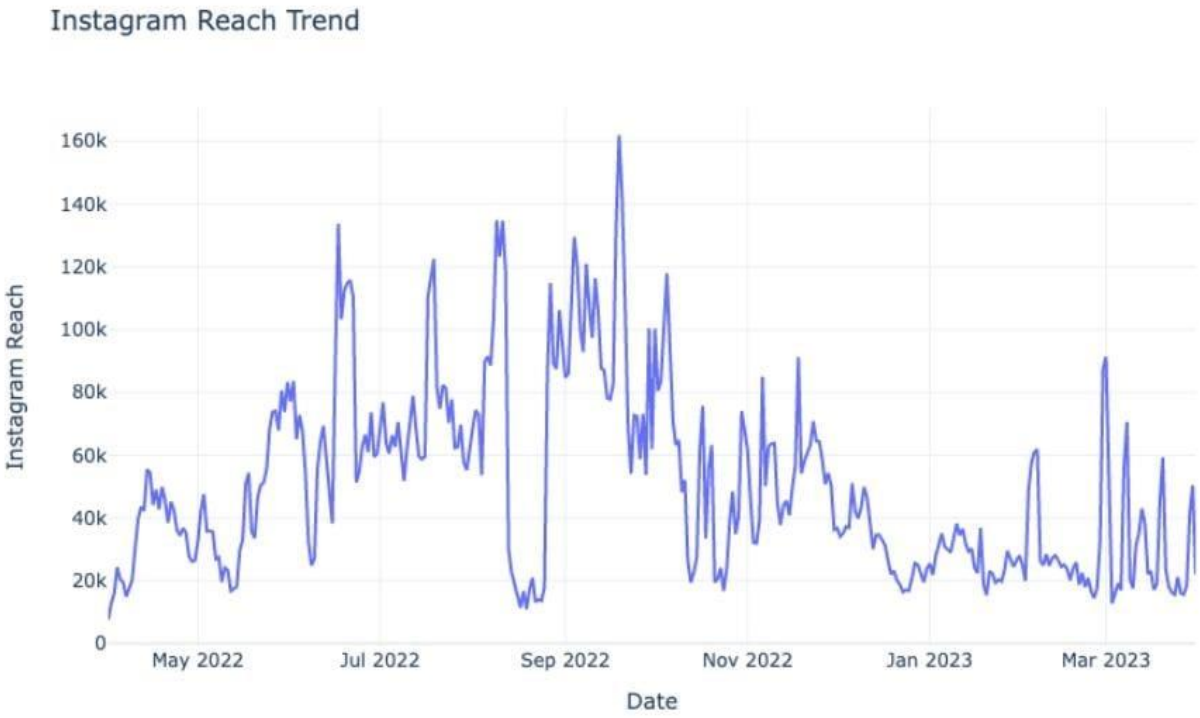
	Date	Instagram reach
0	2022-04-01	7620
1	2022-04-02	12859
2	2022-04-03	16008
3	2022-04-04	24349
4	2022-04-05	20532

Analyzing Reach:

Let’s analyze the trend of Instagram reach over time using a line chart:

```
fig = go.Figure()
fig.add_trace(go.Scatter(x=data['Date'],
   y=data['Instagram reach'],
   mode='lines', name='Instagram reach'))
fig.update_layout(title='Instagram Reach Trend', xaxis_title='Date',
                  yaxis_title='Instagram Reach')
fig.show()
```

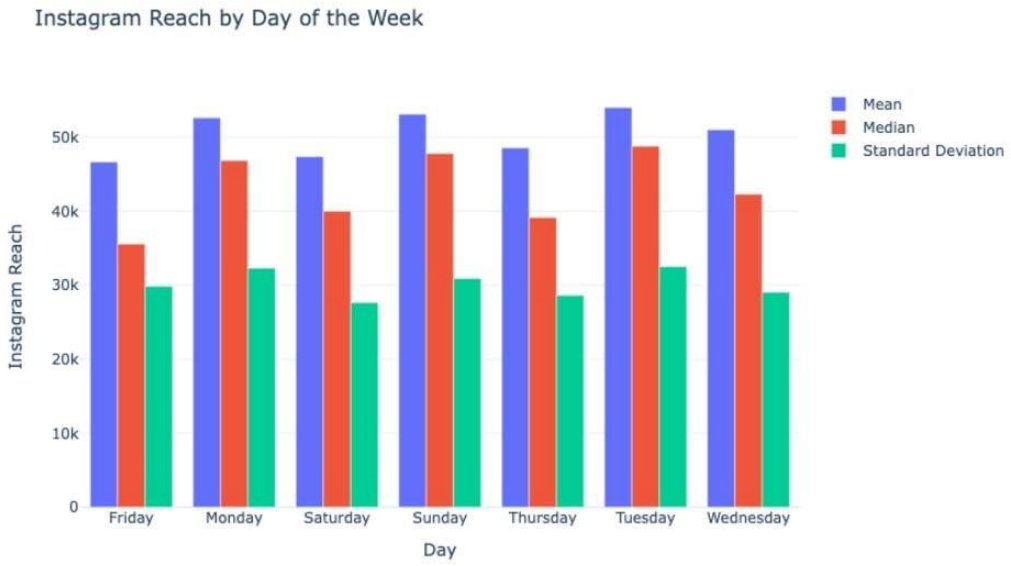
Output:



Now let's analyze Instagram reach for each day using a bar chart

```
fig = go.Figure()
fig.add_trace(go.Bar(x=data['Date'],
                     y=data['Instagram reach'],
                     name='Instagram reach'))
fig.update_layout(title='Instagram Reach by Day',
                  xaxis_title='Date',
                  yaxis_title='Instagram Reach')
fig.show()
```

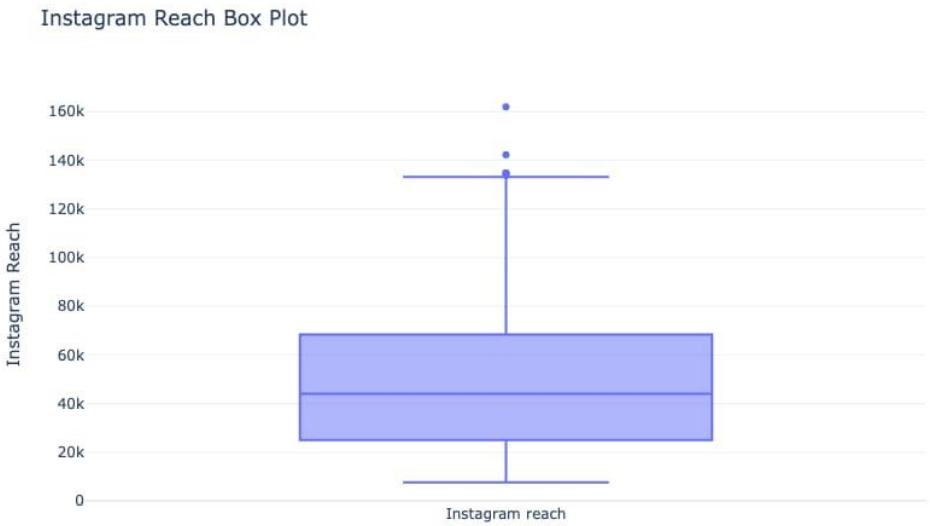
Output:



Now let's analyze the distribution of Instagram reach using a box plot:

```
fig = go.Figure()
fig.add_trace(go.Box(y=data['Instagram reach'],
                     name='Instagram reach'))
fig.update_layout(title='Instagram Reach Box Plot',
                  yaxis_title='Instagram Reach')
fig.show()
```

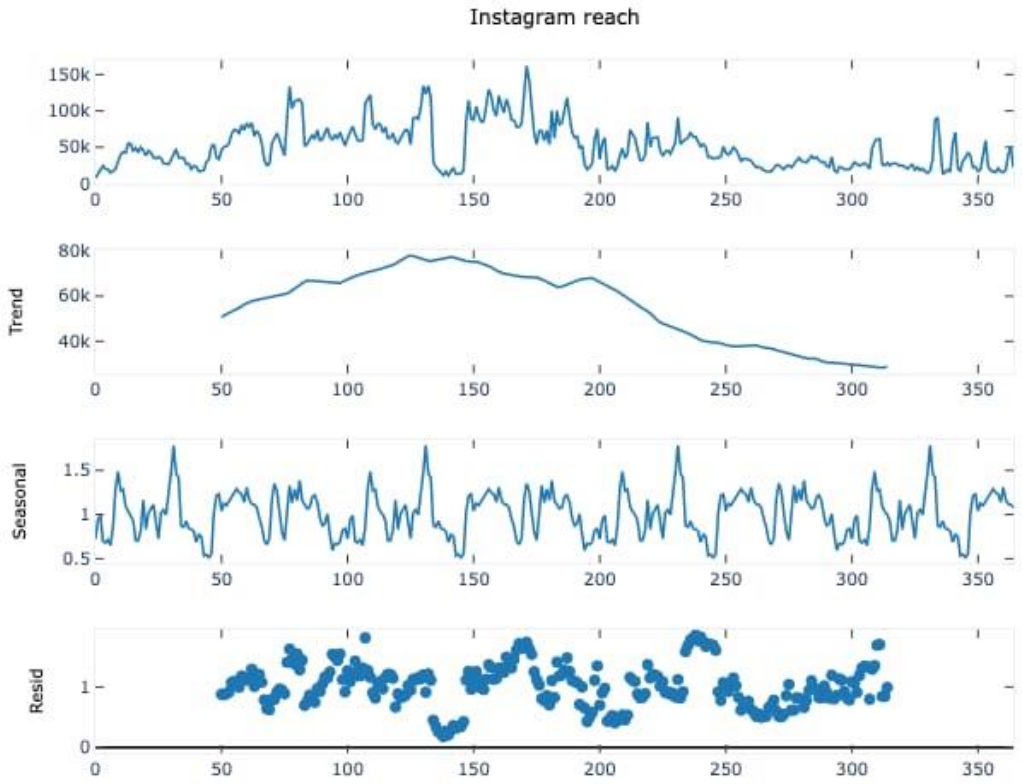
Output:



Let's look at the Trends and Seasonal patterns of Instagram reach::

```
from plotly.tools import mpl_to_plotly
import matplotlib.pyplot as plt
from statsmodels.tsa.seasonal import seasonal_decompose

data = data[["Date", "Instagram reach"]]

result = seasonal_decompose(data['Instagram reach'],
                            model='multiplicative',
                            period=100)

fig = plt.figure()
fig = result.plot()

fig = mpl_to_plotly(fig)
fig.show()
```

Output:



5.TESTING

5.1. TESTING METHODS

SYSTEM TESTING:

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the

Software system meets its requirements and user expectations and does not fail in an unacceptable

manner. There are various types of test. Each test type addresses a specific testing requirement.

TYPES OF TESTS:

1. UNIT TESTING

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

2. INTEGRATION TESTING

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing.

.

1. FUNCTIONAL TESTING

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected.

Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised.

Systems/Procedures: interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

SYSTEM TEST:

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

WHITE BOX TESTING:

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

BLACK BOX TESTING:

Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box .you cannot "see" into it. The test provides inputs and responds to outputs without considering how the software works.

Test strategy and approach

Field testing will be performed manually and functional tests will be written in detail.

Test objectives

- All field entries must work properly.
- Pages must be activated from the identified link.
- The entry screen, messages and responses must not be delayed.

Features to be tested

- Verify that the entries are of the correct format
- No duplicate entries should be allowed
- All links should take the user to the correct page.

Integration Testing

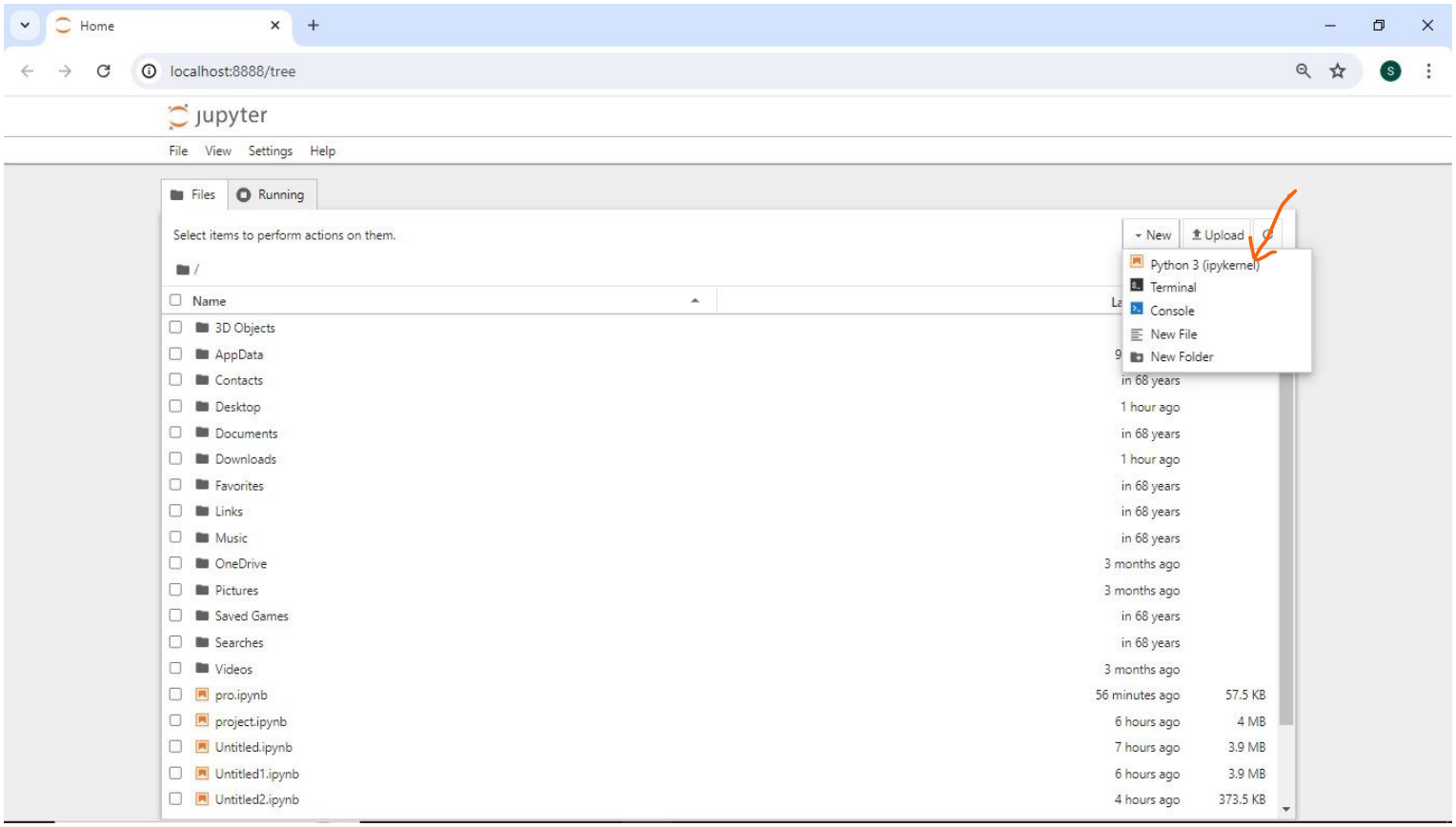
Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects.

Acceptance Testing

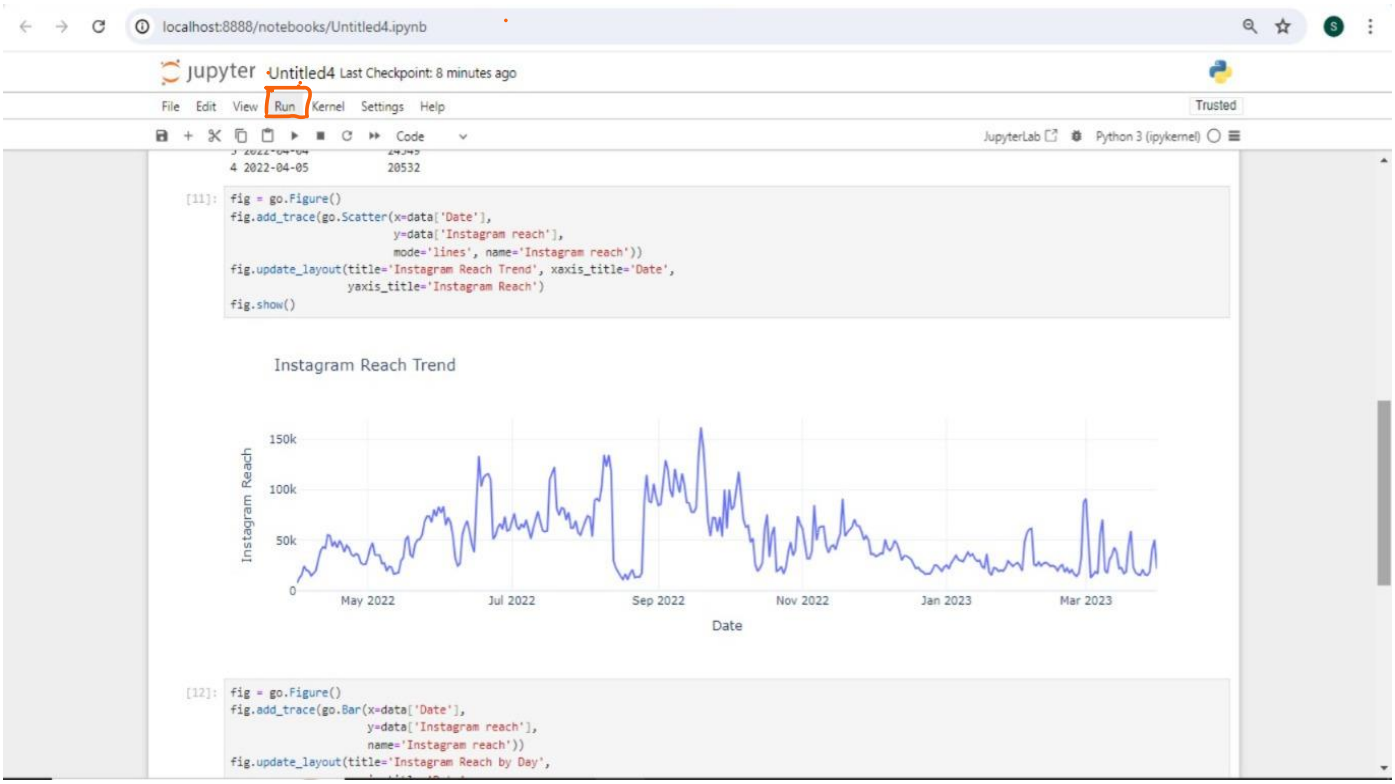
User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the sytem meets the functional requirements.

RESULT:

We have to select the **Python 3(py kernel)**



Import the pandas and write the code
Next Click on **RUN:**



CONCLUSION:

Our goal is to help Instagram users—whether they are marketers, small business owners, or individual content creators—make informed decisions that lead to greater success on the platform. By the end of this study, readers will have a clearer understanding of Instagram reach and practical tools to enhance their own social media strategy.

Instagram stands as one of today's most widely embraced social media platforms. Among its diverse user base, individuals leverage Instagram for professional purposes like business promotion, portfolio development, blogging, and content creation. Given its extensive user reach across various niches, Instagram continually evolves to enhance its offerings for content creators and its user community, striving to provide a better experience for all. This article will guide you through the process of conducting an Instagram Reach Analysis using Python. This analysis is aimed at assisting content creators in comprehending how to adjust and adapt to Instagram's ongoing transformations over the long term.

Instagram is one of the most popular social media applications today People using Instagram professionally for promoting their business. So, it's most important to connect with audiences. So, that we need to know about reach of own account in order to get success. By collecting the dataset of our account which helps us to predict which area has high interaction, so that we can make change accordingly. When we examine hashtag impressions, we realize that not all posts are equally discoverable through hashtags. However, hashtags offer a valuable avenue for reaching new users who share an interest in the content we create.

REFERENCES

1. Ching Wu Chu and Guoqiang Peter Zhang, “A comparative study of linear and nonlinear models for aggregate retails sales forecasting”, Int. Journal Production Economics, vol. 86, pp. 217- 231, 2003.

2. Wang, Haoxiang. "Sustainable development and management in consumer electronics using soft computation." Journal of Soft Computing Paradigm (JSCP) 1, no. 01 (2019): 56.- 2.

3. Suma, V., and Shavige Malleshwara Hills. "Data Mining based Prediction of Demand in Indian Market for Refurbished Electronics." Journal of Soft Computing Paradigm (JSCP) 2, no. 02 (2020): 101- 110

4. Giuseppe Nunnari, Valeria Nunnari, “Forecasting Monthly Sales Retail Time Series: A Case Study”, Proc. of IEEE Conf. on Business Informatics (CBI), July 2017.

5. https://halobi.com/blog/sales-forecasting-five-uses/.

6. Zone-Ching Lin, Wen-Jang Wu, “Multiple Linear Regression Analysis of the Overlay Accuracy Model Zone”, IEEE Trans. On Semiconductor Manufacturing, vol. 12, no. 2, pp. 229 – 237, May 1999.

7. O. Ajao Isaac, A. Abdullahi Adedeji, I. Raji Ismail, “Polynomial Regression Model of Making Cost Prediction In Mixed Cost Analysis”, Int. Journal on Mathematical Theory and Modeling, vol. 2, no. 2, pp. 14 – 23, 2012.

8. C. Saunders, A. Gammerman and V. Vovk, “Ridge Regression Learning Algorithm in Dual Variables”, Proc. of Int. Conf. on Machine Learning, pp. 515 – 521, July 1998.IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 56, NO. 7,