

DATATHON ON A PAGE

Predicting and automating the top 10 stock performers across US stock markets

Powered By:

THE SHAILY SHOW

HOSTED BY

Shailendra Kumar, Global Analytics Leader

Click for more details: https://bit.ly/YT_SKS

Business Problem

Predict top 10 stocks across the entire US market (NYSE, NASDAQ, etc.)
automated each week using the daily stock market data as the base data:

- ✓ Predicted Value and the Percentage Increase
- ✓ Explain the reasons behind the Prediction and the Increase

The process needs to be automated end to end and run each week on Sunday



Base Data: Daily historical Data is freely available and can be downloaded from internet through the websites like Yahoo Finance, Google Finance, Stooq
Other Freely available Data: Social Media, Newsfeeds, Weather etc..... **Be Creative**



Approach



Statistical Modelling



Model Accuracy



Presentation

GUIDELINES

Things to keep in mind

Powered By:

THE SHAILY SHOW

HOSTED BY

Shailendra Kumar, Global Analytics Leader

Click for more details: https://bit.ly/YT_SKS



Technology

- Language to be used: python
- Only use open source and freely available models , frameworks and software(s)
- Utilization of other structured and unstructured data sources is encouraged
- Consider and explain data acquisition techniques, data curation, data storage for quick retrieval, test data vs training data.



Process

- Document the analysis, approaches considered and why specific approaches were used as compared to others
- Explain which models were considered , why each model was used as compared to others - SWOT analysis
- Document and annotate code



Results

- Use open source or freely available data visualization tools
- Present intermediate results for tracking and training data and explain how they contributed to the final output
- Compare predicted and actual stock values – measure model accuracy, train the model to improve accuracy

EVALUATION CRITERIA

Measure the success

Powered By:

THE SHAILY SHOW

HOSTED BY

Shailendra Kumar, Global Analytics Leader

Click for more details: https://bit.ly/YT_SKS

Data Visualization

- Visualization of various patterns
- Presentation of data in intuitive and easy-to-understand visualizations

Model Accuracy

- Variance between predicted and actual stock values
- Which variables have maximum and least impact on model accuracy



Solution and Code Documentation

- Overall solution documentation
- Code documentation

Model Explainability

- As per selected model what data features are most important ?
- For every prediction from the model, how did each feature in the data affect that particular prediction
- What interactions between features have the biggest effects on the model's predictions

REWARD

Celebrate the success - brilliant teams / individuals have the opportunity to get internship offers

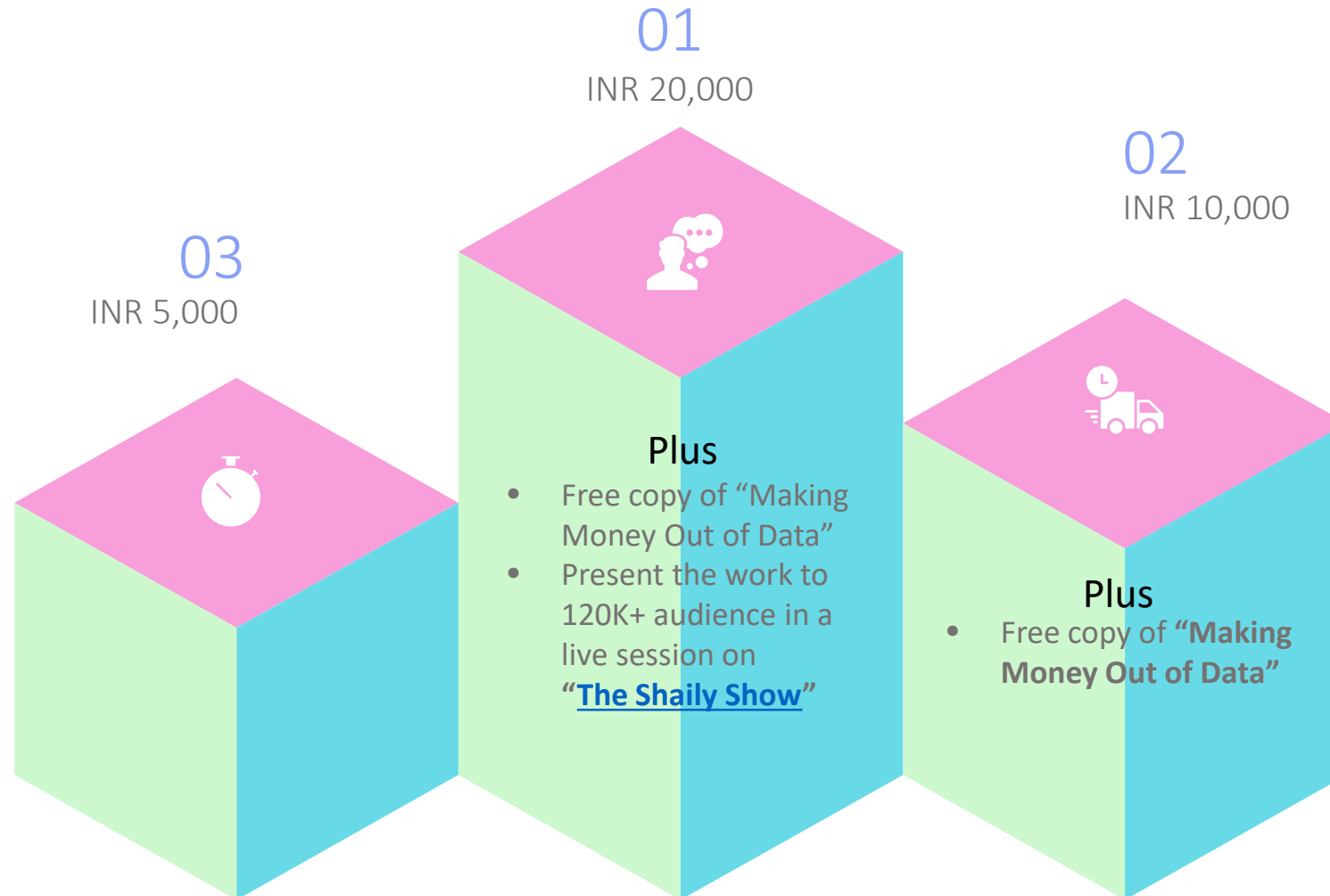
Powered By:

THE SHAILY SHOW

HOSTED BY

Shailendra Kumar, Global Analytics Leader

Click for more details: https://bit.ly/YT_SKS



MEET THE JUDGES

Top Analytics Leaders will be judging the Datathon.

Powered By:

THE SHAILY SHOW

HOSTED BY

Shailendra Kumar, Global Analytics Leader

Click for more details: https://bit.ly/YT_SKS



Shailendra Kumar

Global Analytics Leader
VP & Chief Evangelist, SAP

Shailendra is the author of
the book **"Making Money
Out of Data"**



Sanjay Sharma

Managing Director, Accenture AI,
Strategy & Consulting, India

Sanjay leads the Analytics
and AI practice in India



Ashish Sharma

ANZ Data Analytics Leader
Ex-Analytics Lead, ANZ, Capgemini

Ashish has built and led
Advanced Analytics and Data
Science Consulting teams
across ANZ



Data Science is a creative process of using different data sources to solve a business problem.

SHAILENDRA KUMAR