Arpan Mishra

✉ mishraarpan6@gmail.com 📱 09958325792 📍 New Delhi, India 🌐 arpan-mishra.github.io/

EDUCATION

B.Sc(H) Statistics
Kirori Mal College, University Of Delhi

07/2018 - Present   *CGPA 8.35*

Machine Learning and Statistical Analysis
World Quant University

03/2020 - 06/2020   *passed with honours*

Deep Learning Specialization
deeplearning.ai

05/2020 - 05/2020

WORK EXPERIENCE

Data Science Associate
ZS Associates

11/2021 - Present

Research Intern - Statistics
Inria - Modal, University of Lille, France

06/2021 - 09/2021

Achievements/Tasks

- Worked with medical data for mental health patients with history of suicide attempts.
- The objective was to model the recurrence of a suicide attempt from demographic as well as medical survey data by VigilanS using parametric as well as non parametric statistical methods.
- We also conduct spatial analysis of the patients and use geostatistical techniques to include the effect of spatial autocorrelation into the model.
- Used tools: R, Python, GLM, GAM, Spatial Probit etc.

Machine Learning Engineer (Part Time)
Omdena

08/2020 - Present

Achievements/Tasks

- Worked with satellite imagery and survey data from Census and DHS.
- The objective was use satellite imagery to predict those socio-economic indicators of regions in India, which can act as a proxy for economic well-being.
- Used tools like google earth engine python api, rasterio, qgis, fast ai, pytorch, keras, AWS etc.

Data Analyst Intern
Mindler

06/2019 - 07/2019

Achievements/Tasks

- Worked with the sales and advertisement teams in order to draw insights from the website traffic data as well as the data created by the sales team.
- Analysed various metrics such as User Traffic, Bounce Rate, Exit Percentage etc. to help the organisation plan their ad campaigns efficiently.
- Tools used: R, MS Excel, Google Analytics

SKILLS

`Python` `R` `SQL` `Statistics` `Data Science` `Machine Learning` `Deep Learning` `NLP` `Pytorch` `Keras` `Fast AI` `AWS`

PERSONAL PROJECTS

Rossmann Sales Prediction (10/2020 - 10/2020) ↗
- Created a tool to predict the daily sales of any store of the Rossmann Drug Store Chain which is the 2nd largest drug store chain in Germany.
- Created features referring the active promos run by the store, competitor distance, competitor time of opening, demarking the start of year/quarter etc. and used machine learning algorithms like Random Forest, XGBoost and Neural Networks.

Global Suicide Analysis 1985-2015 (04/2020 - 04/2020) ↗
- Analyzed the global suicide data for 90+ countries from the year 1985 - 2015 in R.
- Used linear models to test the hypothesis of linear relationship between the suicides per 100k population and various features like population, age, sex, GDP per capita etc.
- Various statistical techniques and data visualizations like barplots, boxplots and density curves were used to explain the data.

Sentiment-Extraction-using-Bert (06/2020 - 06/2020) ↗
- Using Bert to detect the sentiment of a given text and extract the words which contain the detected sentiment.
- Created a Flask API endpoint which was hosted on a local web server.
- Pytorch and Hugging Face library were the core technologies used.

Anime-Generation-using-Deep-Learning (07/2020 - 07/2020) ↗
- The aim of the project was to see how far technology has come in just a few years when it comes language models.
- I used two techniques, LSTMs and then a fine tuned GPT2 for generating text and the results were astounding!
- Pytorch and Hugging Face were the core components.

ACHIEVEMENTS

AI Crowd Blitz Hackathon - Rank 9/348 (05/2020 - 05/2020)

IIT Guwahati HR Attrition Hackathon - Rank 20/900 (06/2020 - 06/2020)

LANGUAGES

English
Full Professional Proficiency

Hindi
Full Professional Proficiency

Oriya
Full Professional Proficiency

INTERESTS

`Chess` `Playing the Ukulele` `Graphic Designing` `Table Tennis`