Ravinder Saluja

DATA SCIENTIST

Mumbai Central, Mumbai, Maharashtra, India, Pin Code - 400008.

linkedin.com/in/ravindersaluja | ✉ ravinder.saluja@yahoo.com | github.com/ravindersaluja | ▯ (+91) 9870609759

Experience

Tech Mahindra Business Services   Mumbai, India

ANALYST   May 2019 – Present

- Building predictive models using Machine Learning algorithms for various business projects.
- End-to-end project management, communicating with stakeholders, understanding the business requirements and delivering projects accordingly.
- Proposing solutions and strategies to business challenges and taking suggestions from stakeholders for improving the developed models.
- Identifying valuable data sources for automating various manual tasks with the help of Python and delivering it as a Windows executable file / scheduling the process with batch files.
- Preparing decks and dashboards for better visualization of the data, thus making derivation of insights easier.
- Leading a team, for scraping data from the web and building a dashboard for improving the overall customer experience.

MANAGEMENT TRAINEE   Jan 2019 – Apr 2019

- Worked on building a text classification model – Naïve Bayes using Natural Language Processing techniques and the sklearn module for predicting the contact reason of the customer which helped with efficiently dealing with customers.
- Undertook preprocessing of structured and unstructured data.
- Built dashboards using D3.js for showcasing the key contact center KPIs in an interactive manner.

Edwise International LLP   Mumbai, India

FACULTY   May 2017 – Jun 2018

- Taught quantitative analysis to students appearing for international standardized tests like GRE, GMAT and SAT in group settings.
- Trained newly recruited faculty members for different branches.

Education

S.P. Jain School of Global Management   Mumbai, India
Professional Certificate Program in Big Data and Analytics.   Class of 2018
Watumull Institute of Electronics Engineering and Computer Technology, University of Mumbai   Mumbai, India
B.E. Electronics and Telecommunication.   Class of 2016

Projects

Role: Analyst, Tech Mahindra Business Services

Current Project: MEDALLIA-BOLDCHAT THEME ALIGNMENT, Python (pandas, nltk, sklearn - Naïve Bayes)

- The objective was to map the themes present on the Medallia's portal across the survey verbatim data and align the same with BoldChat's survey data. The challenge was there were multiple themes assigned to surveys. So, this was a Multi-Label Multi-Class classification problem.
- The objective was to prepare multiple classification models that would categorize the survey verbatim and those models trained, would help predict multiple themes for a single survey which would then be mapped against BoldChat's surveys.
- Approach: Clean the text data using various methods of the NLTK/textblob library, create custom methods for correcting spelling mistakes, preparing word list dictionary and n-grams, and finally creating a Document Term Matrix from the corpus with the help of TfidfVectorizer which was passed to the Naïve Bayes classification model.

CHANGE OF TARIFF (UPGRADES), Python (pandas, sklearn - kNN)

- The objective is to identify customers who are likely to upgrade their tariff by creating a predictive model, trained on the customer demographics, usage, history of interactions and other derived attributes and also providing the probability rate and likelihood of the customer opting for upgrade.
- The customers thus identified will be converted through various campaigns which will help with the uplift in margin for the organization.
- Approach: Building project specific data mart, EDA, under-sampling data, developing a pipeline with sklearn for preprocessing/data wrangling and model building along with cross-validation, model evaluation, tweaking, deploying.

SPEECH ANALYTICS, Python (pydub, ffmpeg, SpeechRecognition)

- The aim of this project was to transcribe the speech of the advisor and the customer into text and then further do the speaker diarization for the same.
- Approach: Converted the '.au' audio files to '.wav' files and split the large audio files into smaller chunks so as to pass it to Google's method of the Recognizer() class from speech_recognition module for conversion of speech to text.
- This helped the Quality Monitoring Team in making the audit process faster and thus requiring a smaller number of FTE's for the given task.

SOCIAL MEDIA ANALYTICS, Python (pyquery, requests, BeautifulSoup, textblob, TensorFlow, Transformers)

- Developed scrapers for social media websites like Twitter and Trustpilot for scraping the tweets and reviews about the client and the competitors, developed sentiments analysis modules for the same as well as Word Clouds for better visualization.
- Deployed this on the AWS EC2 Instance so that the data could be fetched at a scheduled time and the dashboard hosted on the internal server can be updated accordingly.

NEXT BEST ACTION, Python (pandas, nltk, textblob, sklearn - Naïve Bayes)

- The objective was to prepare an NLP based classification model that would categorise the type of text data received on real time basis and give the advisor suggestions as to what next steps should be taken.
- Approach: Clean the text data using various methods of the NLTK library as well as create custom methods for correcting spelling mistakes, preparing word list dictionary and n-grams, and finally creating a Document Term Matrix from the corpus with the help of CountVectorizer which was passed to the Naïve Bayes classification model.

UK MOBILE MARKET WATCHDOG, Python (selenium, AWS EC2 Instance)

- Developed scrapers for various telecommunication websites in the UK with the help of selenium and extracted fields related to price plans which helped with the tariff comparison and deployed these scrapers on AWS EC2 Instance so that the data could be fetched at a scheduled time and the dashboard hosted on the internal server can be updated accordingly.

Role: Management Trainee, Tech Mahindra Business

CUSTOMER PHONE BILL - PDF SCRAPING, Python (PyPDF2, re, os, zipfile)

- The objective was to scrape the phone billing information of customers from unstructured pdf files stored in zip archives, transform it into a structured form and load the information thus scraped to the SQL Server database.

ATTRITION PROJECT ETL, Python (sqlalchemy, pyodbc, win32com)

- The aim of this project was to build an entire pipeline for automating the process of building a data mart for the Employee Attrition Project right from extracting the spreadsheets from Microsoft Outlook to transforming it into the required format and finally loading it to the SQL Server database.

Tools/Frameworks

Data Preprocessing: Pandas, Numpy.
Machine/Deep Learning Implementation Frameworks: Scikit-Learn, Keras, TensorFlow.
Machine/Deep Learning Algorithms: Logistic Regression, Naïve Bayes, kNN, XGBoost, CatBoost, LGBM, ANN (Keras | TensorFlow backend), CNN, RNN.
Natural Language Processing: NLTK, Textblob, Gensim.
Development: Selenium, Requests, BeautifulSoup, Lackey, PyInstaller, Microsoft SQL Server, Git.
Data Visualization: Pandas-profiling, Tableau, D3.js, Matplotlib, Seaborn, Bokeh, Plotly.
Cloud Services: AWS EC2 Instance, Google colaboratory, IBM Watson studio, Kaggle kernel.
Integrated Development Environments/Text Editors: Spyder, Jupyter Notebook, SQL Server Management Studio (SSMS), Sublime Text.

Community _____

Organisation/Groups

- Member at Google Developers Group.
- Member at PyData – Mumbai.
- Member at AIDL (Artificial Intelligence & Deep Learning)

Additional Certifications _____

- IBM AI Engineering Specialization   Credential ID: 9FYWT34G9QJT
  - Machine Learning with Python   Credential ID: HMZ5P46RKLXT
  - Scalable Machine Learning on Big Data using Apache Spark   Credential ID: VFY97GAB56C6
  - Introduction to Deep Learning and Neural Networks with Keras   Credential ID: DDUX7QERNGBM
  - Deep Neural Networks with PyTorch   Credential ID: EPAF7H38KSET
  - Building Deep Learning models with TensorFlow   Credential ID: 3TXZKQLR6HMU
  - AI Capstone Project with Deep Learning   Credential ID: NSR88ECVB66A

Achievements _____

- Received the "Ace of Spades" Award as an Analyst from Tech Mahindra Business Services for delivering quality projects using Machine Learning Algorithms for Q4 2020.
- Graded "A" in overall performance and project delivery as a Management Trainee at Tech Mahindra Business Services.

Languages _____

English, Hindi.