Yuvraj Sharma

PERSONAL DATA

Address:	**19.3, LIV Student Housing, Church St, Dublin 07**
Phone:	**+353 899561681, +9198999268615**
Website:	**https://yvrjsharma.github.io/**
Email:	**yuvrajsharma.engg@gmail.com**
Linkedin:	**https://www.linkedin.com/in/yuvraj-sharma-a7154628/**
Status:	**Ireland - Stamp-1g	Masters at TU Dublin**

PROFILE

Graduating **Masters' in Data Science** from TU Dublin in Sep2020. Strong 9 years exp. in Data Analytics & Science. Led a team to develop Online Media Management analytics solution with several KPIs for TESCO UK(Dunnhumby). Experience in designing and developing data science/analytics solutions in Python, **Pyspark**, Bigdata, A/B testing & **Computer Vision**. Experience in Dashboard Visualization in **Tableau**, **Python**, and **R**. Worked on Deep Learning projects like **Text Analytics**, **Image Processing**, using NLP, Transfer learning, CNNs, LSTM, OpenCV, Numpy, Pandas, Scikit, Tensorflow Keras & **SQL**.

EDUCATION

Sept 2019 - Ongoing	Master of Science in Computer Science (Specialization Data Science) **Technological University of Dublin, Ireland** Academic Status: Ongoing (2019-2020)
	Award: **First class honours** degree in taught portion
	Modules: **Deep Learning** using Python (Image, Text Analysis) & **Tensorflow2.0 Keras**, Probability and **Statistical Inference**, Data Mining, Working on Data using **R**, **Machine Learning** using Python 3, **Scientific Research** and Literature, Research Design & Proposal writing, **Data Visualization** (**Tableau and R**), and Advanced Databases
	Key Projects: 1. Computer Vision – Improving Transfer Learning in Multispectral image domain like Medical and Satellite imaging. (Final Dissertation Topic) 2. Performed exploratory analysis of Correlation and Difference (t-test, ANOVA *etc.*) and multivariate inferential statistical techniques like Linear, Logistic Regression, and PCA. 3. Performed data scraping, data cleaning, analytics and created static and animated viz. (key libraries used – dplyr, stringr, ggplot, gganimate and plotly) using R and Tableau. 4. Performed systematic and comparative investigation of Deep Learning tasks of Text Analysis and Text Generation using algorithms like RNNs, CNNs, and LSTMs and Tensorflow2.0 Keras.
April 2009	Bachelor of Engineering in Computer Science **Panjab University, Chandigarh, India** Percentage: 2.1

WORK EXPERIENCE – KEY PROJECTS

July 2018 - **July 2019**	**Lead Applied Data Scientist at Dunnhumby, Gurgaon, India** **Tesco UK Analyst Onsite Media Management Team** Lead and developed the design and implementation of an analytics solution for Onsite Media Management team of Tesco UK. Online data had two sources, Adobe-Omniture click-stream data and Google AdSense data. Solution was developed on HDFS/Hadoop distributed cluster and was operated using spark framework in **python** also called **pyspark**. We designed the data flow for entire solution and developed it as a python configurable module. Many KPIs were created for evaluation of Adobe and Google published campaigns separately. Lot of pyspark optimization was done to reduce the processing time and memory overhead on Yarn scheduler.

July 2016 - **June 2018**	**Analyst at Bank of America, Gurgaon, India** **NLP Imperative for BOA: Customer Feedback Analysis using Text Classification** The objective of this project was to provide some insight based on the customer feedback present in free text format with Bank of America call centers. Based on those insights business took decisions to improve customer satisfaction across all the bank products. We performed various data cleaning and data preparation steps to prepare the input data for **SVM** algorithm to perform for sentiment analysis and used **LDA** for topic modelling. Visualizations in **python** and **Tableau** were done as a part of this project, in which we came up with interesting plots like US state-wise customer sentiment and topics distribution. Most *negative* call centres and states were picked up by the business as an outcome. The code was developed in Python 2.2.
Mar 2015 - **June 2016**	**Analyst at Bank of America, Gurgaon, India** **Python analyst: Multiple projects** Developed a solution in Python 2.1 to a fuzzy string-matching or **pattern matching** problem faced by one of the Bank's team. The objective was to identify False Positives inside the database. The algorithm developed around cosine similarity was able to identify the false positives close to 95 percent of times. Created solution for **Predictive Indent** in Python 2.1 for customers calling on the helpline. Solution was able to predict the intent of customers before routing the call to an executive. 7 Use-cases were developed in the pilot phase and were successfully implemented. Built a **metadata repository automation** in Python 2.1 for bank's Hadoop data lake migration. Automation was able to read table details from database and excel files and gather respective metadata from database support pdfs and doc files. Months' work was reduced to a matter of days and it was widely appreciated during the all-hands meetings.
June 2013 - **May 2015**	**Software Engineer at Bank of America, Gurgaon, India** **Multiple projects and Technologies** Developed and maintained **Informatica ETL** mappings and workflows to build and load data in Data Quality repository. Created **SQL** rules for multiple sources like **Teradata**, SQL Server, and SAS datasets as a part of Data Quality execution team.
May 2010 - **May 2013**	**Systems Engineer at Tech Mahindra, Pune, India** **Barclays CRM team -** Worked as a Teradata SQL Developer for the client **CITCO Nextgen Data Conversion –** Developed and maintained SQLs/PLSQLs for feeding into Oracle warehouse builder ETL tool

CERTIFICATES AND SKILLS

Data Science:	**Data Science for Engineers, 2018** (IIT Madras, NPTEL), **Programming, Data Structures and Algorithms Using Python, 2018** (Chennai Mathematical Institute, NPTEL), **Business analytics and data mining Modelling using R, 2018** (IIT Roorkee, NPTEL) Multiple other certifications from Linkedin Learning and Coursera can be seen and verified on my LinkedIn profile.
Tools:	Jupyter Notebook, R Studio, Tableau, SAS Enterprise Miner, Informatica ETL, SQL developer, Oracle
Interests & Hobbies:	Technology, Space research, Cricket, Reading, Programming