Pavan Kumar Kosaraju

Website: https://pavankosaraju.github.io/ LinkedIn: linkedin.com/in/kosarajupavan
Mail: vkosaraj@sfu.ca

Overview

- Interned at ***Royal Bank of Canada*** as a ***Data Scientist*** in ***Global Cybersecurity team***.
- Skilled in ***PySpark***, ***Pandas***, ***NumPy***, ***Keras*** and ***Pytorch***.
- Experienced in ***Supervised learning***, ***Unsupervised learning***, and ***Graph Analytics***.

Technical Skills

Programming Languages	:	Python (NumPy, Pandas, Matplotlib, Pytorch, Keras, NetworkX), R, C# (.NET Framework).
Big Data Tools	:	Spark (PySpark), Hadoop, GraphFrames, and MapReduce.
Databases	:	MySQL, NoSQL (PostgreSQL, Cassandra).
IDEs	:	Jupyter Notebook, PyCharm, Spyder, R-Studio.
Tools	:	Git SCM, Jira (Kanban), Confluence, GCP (Dataflow, Data Fusion, BigQuery), Dataiku DSS and Tableau.

Work Experience

❖ **Royal Bank of Canada** **Toronto, Canada**
Data Scientist (Global Cyber Security - JSOC Analytics) May, 2019 – December,2019

- Optimized access certification for 11K managers across RBC by analysing employee-to-resource mapping data.
- Performed outlier detection to identify employees with abnormal accesses, enabling RBC to stay secure and compliant.
- Developed a POC model for Bottom-up Role Mining using graph analytics and unsupervised clustering techniques.
- Built end-to-end data pipeline for role mining project using DataIku DSS and Hadoop tools like PySpark.
- Implemented custom one-hot encoder and custom matrix multiplication algorithm for large scale matrices in PySpark.
- Documented data dictionary for our project's data source by coordinating weekly meetings with the business team and solution architects, allowing the team for quick look-ups about the data during EDA.
- Received appreciation from the business team for rapid delivery of project reports.

❖ **Tata Consultancy Services** **Pune, India**
Assistant System Engineer June, 2016 – January, 2018

- Performed troubleshooting and resolved issues for client applications via timely bug fixes and efficiently communicating solution protocols.
- Assisted in automating the emailing process for a web application using a Nintex workflow.
- Continuously monitored the overall health of all the servers maintained by client, as well as tracked various server performance metrics like server uptime, average response time of the server, and server memory utilization in real time, minimizing the security issues of the applications.
- Scheduled server- down-time and supervised patch updates for web applications to ensure a smooth upgrade.

Academics

- **MSc in Computer Science - Big Data (CGPA – 3.82)** September, 2018 – April, 2020
  Simon Fraser University, Burnaby, Canada.
- **Bachelor of Technology in Computer Science (CGPA – 3.7)** June, 2012 – April, 2016
  GITAM University, Visakhapatnam, India.

Projects

- Semantic Search for speeches in Audio (NLP, Embeddings, Keras, Pytorch) – April, 2020. [GitHub]
- Vancouver Housing Market Analysis (Recommendation systems, Time series analysis, Flask) – April, 2019. [GitHub]
- Reinforcing CNNs against adversarial images (Pytorch, NumPy, Deep learning) – December, 2018. [GitHub]