Pranav Kotak

Buffalo, NY | pranavki@buffalo.edu |+1 716 572 7083| https://www.linkedin.com/in/pranavkotak/

EDUCATION

University at Buffalo Aug 2022 – Dec 2023 *Master of Professional Studies in Data Sciences and Applications*
Relevant coursework: Statistics & Probability, Data Mining, Database Management Systems, Data Management and Design
Mukesh Patel School of Technology Management & Engineering, India Jul 2018 - May 2022 *Bachelor of Technology in Data Science* **GPA: 3.96/4.0 Relevant coursework:** Programming in C++, Data Structures and Algorithms, Machine Learning, Deep Learning, ETL, Big Data, Data warehousing, and Mining

TECHNICAL SKILLS

Programming Languages: C, C++, Java, SQL, Python, JavaScript, MATLAB, R
Machine Learning: Scikit-learn, Pandas, OpenCV, TensorFlow, PyTorch
Software & Tools: MS Excel, SAS, Tableau, Google Analytics, Android Studio, MongoDB, PostgreSQL, Airflow, NLTK, PowerBI
Web Technologies: HTML, CSS
Certifications: Machine Learning with Python (IBM

PROFESSIONAL EXPERIENCE

Bangalore, India Nov 2021 – May 2022 *Great Learning – Data Scientist Intern*
- Conducted data analysis on 200 students' data from 5 internal courses in 5 distinct programs to design a new course that includes practical application of the tool Airflow. Research for specific courses included an analysis of tool's existing demand, integration with other tools, and reusability.
- Maintained content for internal documentation and knowledge sharing.
- Contributed to designing 4 case studies and 1 capstone project from start to finish in Airflow, each requiring creation of fresh data.

SIGNZY Technologies, Mumbai, India Apr 2021 – June 2021 *Business Analyst Intern* **(Python (sklearn, pandas, and SciPy for mathematical operations), MySQL)**
- By combining several database tables, one million rows of raw user data were transformed into usable data (data processing).
- 7 Tableau dashboards were created for various scenarios to better analyze user behavior over a brief period.
- Worked in depth with data to get some basic statistics (5-number summary) about it.

ACADEMIC PROJECTS

Text Generation Portal (Python, NLTK, sklearn, HTML, CSS) Jan 2021
- A web application that accepts beginning of a text as input and forecasts its continuation based on the user-specified number of words to predict next.
- Made a customized LSTM model with a ROUGE score of 0.86 and test accuracy of 90%.

Audio Analytics (Python, R, PyPI) Dec 2020
- Developed a Python module that takes an audio file as an input and outputs the emotion in given audio. • Used python libraries to analyze different audio parameters.
- Built a 1D CNN model that performed better than the 2D CNN models already in use and provided an accuracy of 80%.

RESEARCH

Unmasking Privacy Leakage through Android Apps Obscured with Hidden Permissions IEEE - Dec 2021
- 500 Android apps were categorized into low, medium, and high-risk categories using a novel algorithm based on hidden and dangerous permissions.
- To create the risk metric score for apps, a technique based on linear regression was applied.
- Method was verified using K-means clustering to match the app permissions' original data. To correlate the results, 3 separate clusters were created.

Enhancing the Data Mining Tool WEKA IEEE - Oct 2020
- Developed 3 strategies to get around WEKA's main flaws.
- The CSV to XLSB file format conversion has been tested in experiments to validate the methodologies; it is a big change.