Harrison S. Jansma

github.com/harrisonjansma

(210) 251-1512 | harrisonjansma@yahoo.com

EDUCATION

The University of Texas at Dallas (Part-Time)   **Aug 2018 – May 2021**
MS Computer Science, Data Science Track

Baylor University   **Aug 2013 – May 2017**
BBA Business Fellows, Mathematics

TECHNICAL SKILLS

- Highly proficient (2-3 years): Python, Statistical Programming, SQL, Machine Learning Data Visualization, TensorFlow, Unix, Bash, Git
- Proficient (<2 year): Java, Scala, HIVE, Spark, Docker, AWS

WORK EXPERIENCE

Data Process Manager– Capital One   Plano, Texas   **April 2020 – Present**
Responsible for the automation of divisional reporting. Designed dashboards and automated data extraction methods to improve the speed and accuracy of governance reporting.
- Created data extraction solutions with Python that automate monthly reporting to division leadership.

Data Science Intern – Sprint   Overland Park, Kansas   **May 2019 – Nov 2019**
Instrumental to the design, and implementation of machine learning systems that predict logic failures in Sprint's billing systems.
- Explored data lakes and created high performance HIVE and SQL pipelines that added hundreds of new data points for use in production ML systems.
- Prototyped tree-based models to predict billing errors. Presentation of results to regional VP led to a three-month extension of internship.
- Created and taught a weekly, high-level course on ML methods for managers. Resulted in better team communication and the creation of a new analytics project.

Machine Learning Consultant - Upwork   Plano, Texas   **Oct 2018 – Aug 2019**
Natural language processing research with Sociologists at the University of Pennsylvania.
- Designed Python scripts to collect data from millions of social media profiles. Scripts were faster and able to collect data further back in time than the official Twitter API.
- Created large scale data cleaning and processing pipelines for gigabytes of text data.
- Trained a recurrent neural network for emotion-classification of text in TensorFlow.
- Communicated with project leads to create research methods and deadlines.

PERSONAL PROJECTS   Plano, Texas   **Jan 2018 – Mar 2020**
- Implemented ResNet, ResNeXT, MobileNets, and other Conv Nets in TensorFlow.
- Implemented Decision Trees, XGBoost, and Random Forest for text classification.
- Designed and coded a SQL database in Python that supports byte-level compression.
- Scraped 1.2 million articles from Medium.com for data analysis and visualization.
- Designed and trained RNNs, LSTMs, and Attention networks for NLP (PyTorch)
- Created deep learning pipelines with AWS EMR for Spark and Hadoop systems.