Allison Kahn

arkahn@seas.upenn.edu • allisonkahn.com • linkedin.com/in/allison-kahn/ • github.com/AllisonRKahn

EDUCATION

University of Pennsylvania | Philadelphia, PA Expected Graduation: Dec. 2022
M.S.E. Data Science *GPA: 3.95*
Coursework: Applied Machine Learning, Big Data Analytics, Statistics for Data Science, Natural Language Processing, Data Ethics, Data Science for Public Policy, Forecasting for Management, Software Engineering

University of Florida | Gainesville, FL May 2019
B.S. Computer Science *GPA: 3.62*

SKILLS & TECHNICAL TOOLS

Languages: Python, SQL, Spark, SAS, Java, HTML/CSS, JavaScript

Technologies and Tools: Pandas, Power BI, Jupyter, Supervised Machine Learning (regression, classification, time series analysis), Unsupervised Machine Learning (clustering, neural networks, dimensionality reduction), SciKit Learn, Statistical Data Analysis, Git

Certifications: AWS Cloud Practitioner, Microsoft Specialist in Excel & Access

EXPERIENCE

Data & AI Senior Analyst | Accenture Nov. 2020- Aug. 2021
Data & AI Analyst | Accenture Aug. 2019- Nov. 2020

- Developed metrics and SQL scripts to track initiatives worth $23M+ covering financial, clinical, and operational data to support dashboard and power executive-level decision making
- Designed and implemented ETL process and Power BI dashboard to track and visualize KPIs and conducted ad-hoc root cause analysis to explain unexpected results to non-technical leadership
- Collaborated with clients to understand business needs and define problem statements; communicated findings with non-technical business stakeholders through data driven storytelling
- Led effort to create model to represent and quantify inefficiencies in supply chain using image processing and k-Nearest Neighbors (kNN) clustering algorithms in Python, individually identifying $5M in potential savings

Data Scientist Intern | Verizon June 2017- Aug. 2017

- Created a sentiment analysis tool from logs of customer facing live chat data using a Naïve Bayes classifier and natural language processing libraries and techniques in Python
- Contributed to optimization of topic mining algorithm to determine key noun-phrases to summarize chats in Python to identify pain points and high-emotion topics

PROJECTS

Spotify Playlist Continuation | *Python, Spark, Pandas, scikit-learn*

- Explored and evaluated the efficacy of Collaborative vs. Content-Based Filtering predictive models to predict songs a user would enjoy based on songs already in that user's playlist

Book Genre Classification | *Python, NLP, Pandas, scikit-learn, cleaning*

- Built logistic regression model to classify the genre of books based on title; utilized TF-IDF algorithm and extensive preprocessing to clean dataset of 1.7M+ books
- Improved accuracy from 50% to 77% by exploring techniques including over/under sampling, ensemble methods with additional models, tuning hyperparameters and regularization