SPENCER PRENTISS

DATA SCIENTIST

seprentiss@gmail.com 812-572-5678 https://www.linkedin.com/in/spencerprentiss/ https://github.com/Seprentiss

SKILLS

Python (Pandas, Numpy, Skit-Learn), R, Java, SQL (MySQL, SQLite), Excel, SAS, Tableau, Hadoop, Spark

EDUCATION

Purdue University **08/2019-12/2022**

B.S. in Data Science and B.S. in Applied Statistics **GPA: 3.68**

Relevant Coursework: Data Mining and Machine Learning, Information Systems, Theoretical Statistics, Probability, Intro to Time Series, Intro to AI, Large Scale Data Analysis, Applied Regression Analysis

WORK EXPERIENCE

Teaching Assistant - Purdue University-West Lafayette, IN **07/2021-12/2022**

- Taught Java, Python, GitHub, and Unix environments to 200+ students.
- Developed various exercises in Python to teach the fundamentals of programming, data structures, and web scraping.
- Adopted unit testing to reduce grading time of student programming assignments by 2 hours.
- Collaborated with 20 other TA's to answer student issues and prepare resources for study sessions and exam reviews.

Data Analyst Intern **01/2021-05/2021**

- Teamed with a fellow student to assist a high school football team in overcoming a division adversary.
- Employed Monte Carlo simulations in Python to increase expected win performance and expected one-possession game performance by 7% and 13.4% respectively.
- Designed, engineered, and deployed asynchronous and real-time data visualizations in Excel, R, and Tableau.
- Presented findings to coaches and demonstrated how to implement them into game strategy in a 45-minute presentation.

PERSONAL PROJECTS

Purdue Baseball Analytics

- Liaised with Purdue Baseball's Assistant Director of Strength and Conditioning to develop two Tableau dashboards to track and analyze individual and team performance metrics over time.
- Empowered coaches and over 40 athletes to make data-driven decisions by providing actionable insights based on the visualizations.
- Contributed to the optimization of weekly training schedules by providing performance data and insights to the coaching staff, leading to improved athletic performance and team success.

Political Science Research Project

- Spearheaded a 6 person research team to investigate the complexity and polarity of judicial opinions.
- Scraped, cleaned, and analyzed approximately 10GB of data containing opinions from the 12 appellate courts.
- Made use of advanced natural language processing (NLP) techniques, including sentiment analysis and LDA topic modeling, to generate insightful metrics for the complexity, polarity, and subjectivity of court opinions.
- Collaborated with a professor to implement the research findings and methodology into ongoing research efforts.

Loan Default Analyzer

- Developed a Random Forest Classifier model to predict weather someone will default on a loan.
- Conducted feature selection techniques on over 1 million data points to identify the most relevant features for model development, resulting in a more optimized model with improved performance.
- Employed standardization and sampling techniques, such as normalization and random under sampling, to enhance the model's predictive power and reduce bias.

Restaurant Management Web App

- Worked as part of a 5-person team to the develop a restaurant management web application.
- Leveraged the Django framework to integrate a SQLite database for the application.
- Utilized indexing and isolation levels to optimize the performance and consistency of the database.
- Designed and implemented complex triggers and prepared statements to update and manage the application's data.