Sandeep Tiwari

Rohnert Park, CA | (707) 843-1834 | tiwaristiwari97@gmail.com | linkedin.com/in/sandeeptiwari6 | sandeeptiwari6.github.io

EDUCATION

University of Washington **March 2022**
(*Incoming) Master of Science (MS) in Data Science*

University of California, Berkeley **August 2019**
B.A. Data Science and Applied Mathematics

***Relevant Coursework*:** Fundamentals of Data Science, Discrete Math and Probability Theory, Data Structures, Efficient Algorithms and Intractable Problems, Databases, Introduction to Artificial Intelligence, Introduction to Machine Learning, Applied Data Science with Venture Applications, Principles and Techniques of Data Science, Linear Algebra, Economic Statistics and Econometrics

SKILLS

Programming: Python, Java, SQL, MATLAB, Stata, Scheme, HTML, CSS, JavaScript
Software: R, Microsoft Office Suite, Amazon Web Services
General: Unix/Linux, Git

WORK EXPERIENCE

Support Vectors | Fremont, CA **August 2019–*present***
Data Science Intern
- Modeling the spread of COVID-19, using concepts from graph theory and network science. Building large-scale networks with data from thousands of cases to observe which locations are, and predicting which locations will act as hubs for coronavirus

Rebis, Inc. | Palo Alto, CA **January 2019–August 2019**
Data Engineer
- Leveraged social media APIs to parse user data, building profiles of user interests internally within Rebis' platform
- Used data to generate video recommendations based on user interests, through AI-driven mobile recommender system

Snipfeed | Berkeley, CA **September 2018–December 2018**
Software Engineer
- Developed recommendation algorithm for Snipfeed's news engine on various messaging applications. Implemented user-based, item-based, and collaborative filtering recommenders, as well as singular-value decomposition to increase robustness of model

QT Ultrasound, LLC | Novato, CA **June 2018– August 2018**
Machine Learning Research Intern
- Employed deep learning, and computer vision to implement unsupervised learning model, Generative Adversarial Networks, to generate realistic ultrasound images, resulting in sharper imaging and more accurate detection of breast tumors. Currently have a patent pending on the application of this technology.
- Used Pytorch and NVIDIA GPUs.

Fujitsu Network Communications | Sunnyvale, CA **June 2017 –August 2017**
Software Engineering Intern
- Developed full-stack application for integration into Virtuora Network Controller, based upon logging various RPC calls and XML data for networks on scale of thousands of nodes. Also introduced use of ElasticSearch in the network controller and integrated it into the UI
- Wrote Python scripts to scrape key network events from backend logs to display onto custom Flask-based user interface
- Used Virtuora Network Controller, Elastic Search, Python, JavaScript, Flask, ExtJS, XML, JSON, RPCs and REST frameworks

PROJECTS

UW COVID Data Science Hackathon 2020 | Seattle, WA **June 2020**
- Won first place in the inaugural Data Science hackathon hosted by the MSDS program at UW in the data visualization/dashboard category
- Worked with remote team to generate visualizations that compare and analyze how governments of nations across the globe responded to the coronavirus pandemic, and gain insights into how it is spreading, using Python and Tableau

Dynamic Time Series Algorithm for Predicting Long Term Energy Prices | Berkeley, CA **January 2018 – May 2018**
- Built time series model to predict energy price trends by evaluating risks in developing renewable energy infrastructure, emphasizing long-term accuracy and adaptability to significant economic events. Input energy prices over last 15 years with formulated ARIMA model and detailed features
- Through multiple iterations of linear regression and ARIMA, attained highly accurate model, with relatively low mean squared error and an R-squared coefficient of 0.94
- Scraped and pre-processed unstructured government datasets, using Pandas and BeautifulSoup.

Spam-Ham Classification | Berkeley, CA **November 2018**
- Built a text classifier that distinguishes spam emails from non-spam emails, by applying NLP techniques to parse over 10000 emails, attaining an accuracy over 96%, and a precision score of 80%
- Used Python libraries such as scikit-learn for modelling, feature engineering, and document vectorization

LEADERSHIP AND EXTRACURRICULARS

Fundamentals of Data Science Course Staff | Berkeley, CA **August 2017–December 2017**
Teaching Assistant / Undergraduate Student Instructor
- Taught fundamental data science concepts to class of 50 students over the course of a semester in Python. Concepts included hypothesis testing, p-value tests, confidence intervals, and bootstrap sampling
- Tutored undergraduate students in advanced math concepts outside of class, and assisted them with technical issues on homework assignments and projects

Sports Analytics Group at Berkeley | Berkeley, CA **February 2017–December 2017**
Data Analyst
- Used Python and R data analysis libraries to scrape, clean, and process large, unstructured sports data for various research studies