PAVLEEN KAUR

• +1(213)-326-4224 • Los Angeles, California • 19 Nov 1995
• pavleenk@usc.edu • github.com/pavleenkaur

EDUCATION

Master of Science in Data Science, University of Southern California **August 2018 - May 2020**
B.Tech. Information Technology, Indira Gandhi Delhi Technical University, India **August 2014 - May 2018**

TECHNICAL SKILLS

Languages: C, C++, Python, Java, Lua, Spark, Scala,, MySQL, NoSQL, Matlab, D3, Flask, Firebase, JSON, DynamoDB, MongoDB, Algolia, AWS, Wordpress, Magento, HTML, PHP, JavaScript, JQuery, Node.js, Oracle Cloud Analytics, C#
Tools: Visual Studios, Tableau, Trifacta, Weka, Advanced Excel, Adobe Photoshop, Adobe Illustrator, FileZilla, Android Studio
Python : Scikit-learn, Keras, Tensorflow, PyTorch, NumPy, Pandas, Matplotlib, Seaborn, NLTK, Plotly, Scrapy, OpenCV

EXPERIENCE

Koder, Inc - Software Developer Intern, *Irvine, California* **January - March 2020**

- Designed the backend database of Uber as part of the training and established connectivity and interaction through an interface, using SQL Server Management Studio and Microsoft Visual Studio
- Used C, .NET, and Postman for creating and maintaining Get, Put and Post API requests

NEC Laboratories, America - Machine Learning Research Intern, *Princeton, New Jersey* **May - August 2019**

- Facilitated in increasing the efficiency and speed of NEC's Intrusion Detection and Classification System (IDS) by ~80%, and introduced 100% generalization, thereby enabling the re-usability of the system for various applications
- Executed the code conversion of the IDS from the in-house Lua based Machine Learning platform to PyTorch, increasing the accessibility and understandability of the system
- Refined and restructured the incoming data from the Fibre Optics cable, attaining an improved F1 score by ~6%, and accuracy by ~12%, to 94.23%
- Introduced real time detection and classification in the system, with an output latency of 0.5 seconds and TPR of 0.98
- Incorporated the principle behind IDS for the prediction of coefficient of modal decomposition of a light beam, publishing the novel findings in the form of the paper 'Machine-learning orbital angular momentum spectra'

Analytics Vidhya Pvt Ltd - Data Science Intern, *Delhi, India* **May 2017 - July 2017**

- Exploited Python, Tableau and Trifacta for Sales Analysis of a Supermarket in US, and converted the findings into 3 tutorial style articles on Tableau, with a total viewership of 51000 people, increasing the user base of the organization by ~2.7%
- Accelerated the monthly visitors by contributing towards researching machine learning algorithms and generating case studies and skill tests for AV's DataHack portal, that caters to 3.2M monthly visitors

Indira Gandhi Delhi Technical University - Software Development Intern **June 2016 - March 2017**

- Developed a log module using HTML, PHP, JavaScript, CSS and MySQL, to track and log activities of ~300 faculty members, such as registering of courses, log-ins, sign-ups, updating marks, managing coursework; across the IGDTUW Academics and Examination Web Portal
- Supervised the generalization of the portal for use in Admission Services, to maintain records of ~1000 new enrolling students

PUBLICATIONS

- Pavleen Kaur, Giovanni Milione, Eric Cosatto, Philip Ji, NEC Labs. America, Inc., "Machine-learning orbital angular momentum spectra", SPIE Photonics West 2020, San Francisco, United States
- Pavleen Kaur, Payel Ganguly, Saumya Verma, "Bridging the Communication Gap : With Sign Language Conversion", 2018 IEEE/ ACIS 17th International Conference on Computer and Information Science (ICIS), 978-1-5386-5892-5

PROJECTS

Identifying Bias in Airbnb - *Python, Tableau, TensorFlow* **August - December 2019**

- Implemented a Convolution Neural Network to classify host images into different races, with an error rate of 3%
- Visualized and analyzed Airbnb data, and implemented Sentiment Analysis on customer reviews, to identify racial and gender bias
- Used XGBoost to find feature significance and make a fair price predictor that allocated prices to listings based on house features

Art Recommendation - *Oracle Cloud Analytics, Python, Spark, Apriori, MLlib Flask* **January - April 2019**

- Worked on a client project for Kiana Analytics, and built a web platform for a 2 way recommendation system, based on geographic location data as provided by National Art Gallery, Singapore
- Exploited Apriori frequent item set generation to create pairs of visitors and the artworks they visited, further using Jaccard based Locality Sensitive Hashing to find similarity between the various visitors and developing a recommendation system
- Visualized and generated patterns in the footprints of the visitors based on their past visits to the gallery, thereby curating personalized recommendations of artworks to visit, with a standard deviation from the actual path trajectory of ~9%
- Recommended top 5 strategic locations for installation of new artwork to the management, with an RMSE of 0.016

Sign Language Conversion (ASL) - *Python, OpenCV, Tensorflow* **August 2017 - May 2018**

- Carried out Video and Image Processing to extract information from Gestured (Signed) videos of the hearing impaired and utilized spatial and temporal learning through RNN and CNN to predict the signs, with a mean accuracy of 92.3%
- Published the novel implementation and findings in the paper "Bridging the Communication Gap:With Sign Language Conversion"