

Pavan Chhatpar

Atlanta, GA, USA | pavanchhatpar@gmail.com | (857) 930-1785
[linkedin.com/in/pavan-chhatpar](https://www.linkedin.com/in/pavan-chhatpar) | github.com/pavanchhatpar | pavanchhatpar.github.io

EDUCATION

Northeastern University, Boston, MA

Khoury College of Computer Sciences

December 2020

Master of Science in Computer Science, GPA: 4.0/4.0

Related Courses: Machine Learning, Spec. Topics in AI, Algorithms,
Information Retrieval, Parallel Data Processing

University of Mumbai, Mumbai, India

Vivekanand Education Society's Institute of Technology

May 2018

Bachelor of Engineering in Computer Engineering, GPA: 8.99/10.0

Related Courses: Machine Learning, Discrete Math, Analysis of Algorithms,
Calculus, Data Structures, Data Warehouse and Mining

Activities: Technical Officer at Computer Society of India, Executive Head for Praxis

TECHNICAL KNOWLEDGE

Languages: Python, Java, C, C++, Scala, Julia, PHP, Node.js, TypeScript, JavaScript, HTML, CSS

ML & Data Pipeline Tools: TensorFlow, sklearn, transformers, XGBoost, PyTorch, Spark, Airflow, MapReduce, Databricks, Weka

Databases: Dremio, Hive, Vertica, MongoDB, MySQL, MS SQL, Oracle, SQLite

Operating Systems: Windows, Linux (Ubuntu, Mint), Mac OS

WORK EXPERIENCE

Honeywell International Inc, Atlanta, GA

Data Scientist II – Industrial Analytics (HCE)

January 2021 – Present

- Employed **anomaly detection** techniques to reduce fraud in procurement through **stakeholder engagements**
- Implemented a **distributed index** to scale a patent search engine enabling **10x faster** queries using PySpark
- **Minimized costs** and **improved performance** compared to current externally licensed product
- **Mentored** a team of interns on their project, enabling them to be aligned with their goals

Data Science Intern – Industrial Analytics (HCE)

June 2020 – Aug 2020

- Extracted insights from **legal contracts** about measuring supplier performance, using **NLP in Python**
- Optimized data pipelines working on **Azure Data Lake Storage** to reduce latency of feeding data
- Researched various applications of those insights that would help **optimize supplier partnerships**

Wayfair, Boston, MA

May – Dec 2019

Data Science Co-op – B2B/Sales/Service team

- Trained **Survival Analysis Models** on large scale time-series data using **recurrent neural networks in Python**
- Developed data pipelines using **Spark** for data from **Hive**; scheduled daily jobs to run them in **AirFlow**
- Engaged in **stakeholder meetings** to leverage their domain knowledge in **feature engineering**

dotin, Fremont, CA (work from home)

March – June 2018

Software Engineer Intern – Machine Learning

- Developed ML training, testing and predictor modules with pipelining using **Python, Julia, and Java**
- Contributed to maintaining data collection through **Amazon Mechanical Turk**

ACADEMIC PROJECTS & PUBLICATIONS

Deep Question Generation on SQuAD dataset

Master's Project, Northeastern University, Boston, MA

Jan 2020 – Apr 2020

- Developed a deep neural network that generates questions given a paragraph and an answer within it, using TensorFlow 2
- Employing copy mechanism, the generated questions could get answers with an F1 score only **18% lesser** than original ones
- Contributed a generic CopyNet TensorFlow implementation as an **open-source package** via GitHub

The precision of case difficulty and referral decisions: an innovative automated approach

Nair Hospital and Dental College, Mumbai, India

May 2017 – Aug 2019

- Developed an ML solution with a team of dentists to predict difficulty of an Endodontic case prior to treatment using TensorFlow and sklearn with a sensitivity score of **94.96%**
- Published in Clinical Oral Investigations, **Springer**, Aug. 2019

Vehicular Traffic Abatement

Final year Project, University of Mumbai, Mumbai, India

March 2017 – May 2018

- Developed a solution to vehicular traffic using neural networks in a team of four facilitating users with prediction of vehicular traffic based on time and location, with an accuracy of **90.73%**
- Published the project work as two phases in **IEEE**, Nov. 2018 and in **IJRASET** Volume 6, July 2018

INTERESTS/ ACTIVITIES

- Published an android app on Google Play Store, with over 10,000 downloads, to help students in managing attendance
- Won 2nd prize in final year project competition at undergrad college
- Participated in various events on photography, singing, and roller skating