

Atharva Phatak

Thunder Bay, ON, Canada P7E 6M4
(Open to relocating anywhere in Canada)

+1-(807) 355-8351 | athp456@gmail.com | www.linkedin.com/in/atharva-phatak | <https://atharva-phatak.netlify.app/>

PROFILE SUMMARY

- **An inquisitive learner with a zeal to solve challenging problems and drive improvements in projects.**
- 3+ years of programming experience with **C/C++**, **Python**, and **Git** languages.
- **Knowledge of Object-Oriented Design, Data Structures & Algorithms and Software Development.**

EDUCATION

Lakehead University <i>Master of Science, Computer Science. Current GPA: 91.25/100</i> <i>Relevant Coursework: Deep Learning, Machine Learning, Natural Language Processing</i>	May 2021 - April 2023 <i>Thunder Bay, ON, CA</i>
Yeshwantrao Chavan College of Engineering <i>Bachelor of Engineering, Computer Technology. GPA: 8.25 / 10.00</i> <i>Relevant Coursework: Data structures and algorithms, Operating Systems, Software Engineering</i>	April 2016 - Oct 2020 <i>Nagpur, MH, IN</i>

SKILLS

- **Concepts:** Object-oriented Programming (OOP), Data Structures, System Design, Software Engineering
- **Languages:** C/C++, Python, JavaScript, HTML, CSS
- **Database:** SQL, MySQL
- **Operating Systems:** Linux, Windows
- **Data Science:** Machine Learning, Deep Learning, Computer Vision, Natural Language Processing (NLP)
- **Libraries & Frameworks:** Jupyter, NumPy, Pandas, Plotly, React, TensorFlow, PyTorch, scikit-learn
- **Platform & CI/CD Tools:** Jira, Git, Github Actions, DVC

WORK EXPERIENCE

DaTALab Research Assistant	May 2021 – Present <i>Thunder Bay, ON</i>
<ul style="list-style-type: none">• Worked with CRA ITB team to deliver Text-Simplification Pipeline that simplifies CRA documents using NLP based techniques.• Collaborated with an interdisciplinary team of Medical Health Professionals to develop a Text Simplification approach which is simplifies paragraph level biomedical text data.	

PROJECT EXPERIENCE

TorchFlare <i>Python, PyTorch, Numpy, Pandas, PyTest</i>	March 2021 – Present
<ul style="list-style-type: none">• Designed a high-level open source deep learning library using PyTorch as backend to simplify model training and fast experimentation.• Implemented CI/CD procedures using Github Actions for automatic testing and deployment.• Link to the project: https://github.com/Atharva-Phatak/torchflare	
Manga Colorizer <i>Python, PyTorch, Streamlit, BeautifulSoup, Onnx</i>	May 2021 – June 2021
<ul style="list-style-type: none">• Performed novel experiment to convert black & white renders of manga to their colored versions using Pix2Pix model.• Devised a way to improve the color quality of generated images.• Implemented Web Scraper to collect manga images from web sources resulting in a dataset of 3200 high-quality images.• Implemented POC application using Streamlit.• Link to the project: https://github.com/Atharva-Phatak/Manga-Colorizer	
Glassdoor Jobs Analysis <i>Python, Pandas, Matplotlib, Scikit-Learn, BeautifulSoup, eli5</i>	June 2020 – July 2020
<ul style="list-style-type: none">• Implemented a web scraper using BeautifulSoup to scrape 800 job postings related to the title “Data Scientist” from Glassdoor.com• Implemented Random Forest to classify whether a job posting pays more than the median salary.• Achieved AUC-ROC score of 0.78 using Random Forest Algorithm.• Analyzed model’s prediction using machine learning interpretability library eli5.• Link to the project: https://github.com/Atharva-Phatak/Analysing-Glassdoor-Jobs	
Supplements Price Predictor <i>Python, Pandas, Scikit-Learn, BeautifulSoup, seaborn, Matplotlib</i>	May 2020 – June 2020
<ul style="list-style-type: none">• Predicting the prices of fitness supplements using ML models (XgBoost and RandomForests).• Implemented web scraper to collect data about fitness supplements resulting in a dataset with 700 entries and 13 features.• Performed Data Cleaning, Feature Engineering, and Exploratory Data Analysis.• Evaluated model’s performance using metrics like RMSE, MAE, and R2-Score with XgBoost giving the best performance.• Link to the project: https://github.com/Atharva-Phatak/Supplements-Price-Predictor	

ACHIEVEMENTS AND HONORS

- Kaggle Data Science Competitions
 - Team achieved a Global Rank of 76/1621 in Jigsaw Multilingual Toxic Classification using Transformer-based techniques.
 - Team achieved a Global Rank of 221/2227 in Tweet Sentiment Extraction by applying Transformer based techniques.
- Grand Finalist Smart India Hackathon 2018-2019
- Mentored students of Data Science Club and Coding Club, CT department, YCCE.