|  |  |
| --- | --- |
| Experienced Data Scientist, AI, ML and Software Engineer with experiences in Python, AWS, SQL, EDA, ETL, Backend Engineering, and end to end Machine Learning Pipeline which includes Data-Centric AI, ML Modeling, AI Applications and MLOps. | |
| **Experiences** |  |
| ***Lead Data Scientist and Python Engineer, Targetspot*** **|** Manhattan, NY **| Python, AWS, SQL** | **2022 - Present** |
| Working as the primary AI and Data Scientist for AdTech company with Python, S3, SageMaker and SQL   * Using Whisper, NLP models to transcribe, categorize and match podcasts with ads based on listenership * Built a neural network to find podcasts for custom events like Olympics from millions for inventory * Performed ETL, designed and trained AI models from scratch to determine costs of purchasing ad slots |  |
| ***Data Scientist and Software Engineer, Lawfully Inc*** **|** Manhattan, NY **| Python, AWS, SQL, Scala** | **2022** |
| Worked as a data scientist and backend engineer to elucidate USCIS immigration process |  |
| * Scraped, cleaned and De’ID, over 11 million immigration data using AWS S3 and Lambda * Performed EDA and ETL on the scraped data using Python, PySpark, Pandas, Seaborn   + Built model to show case processing times and trends using scraped data [online](https://www.lawfully.com/data/trends/I-765/SRC/2023) * Created scripts for updating and generating new cases using Python, AWS and Zappa * Used scraped data to train Regression and Classification ML models using Keras, TensorFlow, SkLearn and PySpark to predict trends, next steps, and average decision time for immigration cases |  |
| * Wrote the backend code for various new features for over 1 million users |  |
| ***AI Engineering Intern, Huawei*** **|** Ottawa, ON **| Python, TensorFlow, PyTorch, Keras** | **2021 - 2022** |
| * Worked with wireless and traffic data to improve Autonomous driving * Created a simulator in Python to simulate travel pattern of thousands of vehicles in a city * Developed patent pending AI for next gen navigation based on traffic and cellular data * Scraped and cleaned over 32 million data points for training AI models * Built and trained AI models in Python using Keras, TensorFlow, PyTorch and SkLearn * Build Federated Machine Learning models using TensorFlow for thousands of Image and Fintech data |  |
| ***NLP Research Intern, Nuance Communications*** **|** Montreal, QC **| Python, TensorFlow** | **2020** |
| * Implemented new strategies and models to improve Automatic Speech Recognition * Implemented BERT models from hugging face to analyze millions of online product reviews |  |
| ***Machine Learning / Software Engineer, NexJ Health* |** Toronto, ON **|Python, NodeJS, MongoDB** | **2018 - 2019** |
| * Build and trained an NLP classifier model on 8 years of PHI data to categorize patients’ instant messages into different behavioural categories in real time using Python and SQL * Worked on a state-of-the-art health app with NodeJS, Express and MongoDB, used by 100,000+ users * Improved user experience by integrating “login with username” option * Increased user engagement by designing a user points awards system from scratch * Created the mobile app from scratch with Python, SQL on the backend and Flutter, Dart Lang on the frontend |  |
| ***Backend Software Engineer, Montrium*** **|** Montreal, QC **| C#, Python, Java, SQL, React** | **2017 - 2018** |
| * Worked on the backend for a new platform for pharmaceutical study and trial management system, used by over a million employees at over 100 pharmaceutical companies worldwide * Created APIs and primary functions using .NET Core, C#, and Cosmos GraphDB * Developed a Recurrent Neural Network using Keras and TensorFlow, trained on millions of data scraped from multiple sources to translate company’s products to several languages * Built Regression and Classification ML Models from scratch and trained on 100,000+ industry provided data to predict next steps, chances of approval and length of current stage for client’s products * Led a team of five to fully automate unit and system testing using Selenium, Java, and PowerShell to reduce software testing phase from days to mere minutes |  |
| **Education** |  |
| *Fourth Brain Data Science Bootcamp* **| New York, NY** | **2022 - 2023** |
| Machine Learning and AI: Data-Centric AI, ML Modeling, AI Applications and MLOps |  |
| *University of Waterloo* **| Waterloo, ON** | **2019 - 2022** |
| Master of Science **| Computer Science, AI, Machine Learning and Computer Vision**  TA: AI/ML, Functional, C and Concurrent Programming, Data Structures and Operating Systems |  |
| *McGill University* **| Montreal, QC** | **2012 - 2017** |
| Bachelor of Engineering**| Honors Electrical Engineering**, Minor: Software Engineering |  |
| **Thesis** |  |
| Graduate **|** ***Semantic Segmentation*****|** University of Waterloo **|** **Python, PyTorch, Tensorflow** | **2020 - 2022** |
| * Using image level data and approximate class sizes to improve accuracy of Weakly Supervised Semantic Segmentation. Thesis title: [Volumetric Weak Supervision for Semantic Segmentation](https://uwspace.uwaterloo.ca/handle/10012/18321) * Improved the accuracy by over 6% mean Intersection over Union (mIoU) |  |
| Undergraduate **|** ***Image Captioning*****|** McGill University **|** **Python, TensorFlow, OpenCV** | **2016 - 2017** |
| * Used over 100,000 images to train CNN and RNN from TensorFlow to generate image captions with text descriptions in less than 5 seconds using Python, OpenCV and MATLAB for the [Autour](https://autour.mcgill.ca/en/) app |  |
| **Projects** |  |
| *Applied Machine Learning and Artificial Intelligence* **|** **Python** | **2017 - Present** |
| * Built a ML model to detect wake words for Spotify. Wake Word: **Hello Spotify**. [Demo](https://youtu.be/fbmesQSRYfo) * Built a ML model to detect 50 different types of exercise activities from videos * Created a CNN for unsupervised single image depth prediction and a RNN for speech recognition * Created an end-to-end NLP pipeline to classify TV shows and movies * Programmed several different optimization methods, including Gradient and Coordinate Descent, ALM and ADMM, Graph Clustering, noise reduction and several others |  |
| *Various Java, C, Node, React, Angular projects* **| Java, C, NodeJS, ReactJS, AngularJS, MongoDB, SQL** | **2014 - Present** |

\* >= 4 stars -> Advanced

== 3 stars -> Intermediate

<= 2 stars -> Beginner

**Sharhad Bashar**

+1 (646) 683-8387 | [sharhadbashar.com](http://sharhadbashar.com/) | [sharhad.bashar@uwaterloo.ca](mailto:sharhad.bashar@uwaterloo.ca) | [SharhadBashar](https://github.com/SharhadBashar) [|](https://www.linkedin.com/in/sharhad-bashar/)  [Sharhad Bashar](https://www.linkedin.com/in/sharhad-bashar/)



**Honors and Awards**

Mathematics Masters Full Scholarship

**University of Waterloo**

Honors List

Entrance Scholarship

Clifford Knowles Bursary

George Duggan Bursary

**McGill University**

NSERC-USRA

Deans Award

**UWO**

2019 - 2022

2014 - 2017

2012 - 2017

2016

2015

2015

2013

**Languages:**

* **Python**
* **Java**
* **C**
* **Scala**
* **C#**
* **Dart Lang**
* **Javascript**
* **HTML/CSS**
* **MATLAB**

**Libraries:**

* **TensorFlow**
* **PyTorch**
* **OpenCV**
* **SciPy**
* **Scikit-Learn**
* **Keras**
* **Theano**
* **Pandas**
* **AWS**
* **PySpark**

**Frameworks:**

* **ASP.NET Core**
* **Node**
* **DataBricks**
* **React**
* **Selenium**
* **Flutter**

**Database:**

* **PostgreSQL**
* **MongoDB**

**Scripting:**

* **Bash**
* **Git**
* **PowerShell**

å

**Skills\***