

# Sarbottam Thapa Magar

Halifax, NS | 902-8179421 | [Sarbottam.Thapa.Magar@gmail.com](mailto:Sarbottam.Thapa.Magar@gmail.com) | [www.sarbottamthapamagar.com](http://www.sarbottamthapamagar.com)

## Experience

### Data Science Fellow ([http://bit.ly/vehicle\\_prj\\_live](http://bit.ly/vehicle_prj_live))

Dec 2020 – present

SharpestMinds

Built a Full-Stack Data Science product that can predict a used vehicle's price as point estimate as well as range based on its characteristics.

- Worked with scraped dataset from Kaggle containing "458,213 raw data points".
- Built Gradient Boosting Tree based model with a point estimation accuracy 89.52% and an MAE 2,527.49 dollars.
- Built Serverless RESTful API for serving the model with AWS.
- Collaborating through a mentorship with a Senior Machine Learning Engineer at DTN.

### Research Intern

May 2019 – Oct 2019

GreyLit Inc.

Halifax, NS, Canada

Built multimedia search engine using AWS Serverless architecture.

- Developed front-end as well as back-end with serverless architecture and performed quality assurance.
- Built a microservices with lambda to execute the pipeline for text extraction from different multimedia and document files using AWS Machine Learning services (Textract, Transcription, Transcribe) which uses Comprehend & Elasticsearch to index the documents stored in S3 Bucket.
- Tested the whole system with around 1GB of data consisting of different documents (images, docs, pdfs) and multimedia files (video & audio).

### Junior Android Developer

May 2016 – March 2018

Mero Design Pvt. Ltd.

Kathmandu, Nepal

Collaborate with other domain developers to analyze, design, and implement the software products according to client specification.

- Worked on 5 different major projects teaming up in a group consisting 3-4 other domain experts depending on requirements.
- Built Android Mobile Applications using Fragment API, RESTful Web Services, and other third-party libraries using Java MVC.
- Code review, testing, and documentation.

## Education

### MSc. In Computing and Data Analytics

Aug 2018 – May 2020

Saint Mary's University

Halifax, Nova Scotia, Canada

**Relevant Coursework:** Statistics & its Application in Business, Business Intelligence & Data Visualization, Data & Text Mining

### Bachelor's in Computer Engineering

Dec 2011 – Dec 2015

Tribhuvan University

Kathmandu, Nepal

**Relevant Coursework:** Object oriented programming with C++, Data Structure and Algorithms

## Projects

### House price prediction ([http://bit.ly/house\\_project\\_live](http://bit.ly/house_project_live))

Used Random Forest Regressor to predict the price of the house based on important characteristics of house.

- Full-stack data science project involving Exploratory Analysis, Model Building, API Development and Deploying the model in production.
- Built Machine Learning Pipeline using scikit-learn pipeline architecture for automating the machine learning workflow.
- Model was served through RESTful API using Flask and later deployed as Web App in PythonAnywhere server.

### Popular Data Science Tags ([http://bit.ly/pplr\\_tags](http://bit.ly/pplr_tags))

Analyzed data from StackExchange public database to understand about the popular tags being used to post questions in the Data Science community.

- Found out Python tags was 11.17% more used than R in the community.
- Used Apriori algorithm to generate rules for tags used together in the posts.

## Profile

LinkedIn:

[www.linkedin.com/in/sarbotttam](http://www.linkedin.com/in/sarbotttam)

GitHub

<https://github.com/Mattobad/>

Tableau Profile:

[http://bit.ly/STM\\_tp](http://bit.ly/STM_tp)

## Technical Skills

### Programming Languages:

Python | Java | JavaScript

### Data Science:

**Data Analysis** (NumPy, Pandas)

**Data Visualization** (Matplotlib, Seaborn, Plotly)

**BI** (Tableau)

**Supervised learning** (Linear Regression, Logistic Regression, Random Forest etc.)

**Unsupervised Learning** (Clustering, K-means)

**Deep Learning** (TensorFlow-Keras, Image Classification)

**Model Deployment** (Flask-API, Serverless API)

**Other** (Structured, Semi-structured & Unstructured data types, Time Series Analysis, Sentiment Analysis)

### Cloud:

**AWS** (Serverless, S3, Lambda, Elasticsearch, Comprehend, AWS Machine Learning Services)

### Others:

SQL | Bash | Docker | Git

### Operating System:

Windows | Linux | Unix

## Awards & Achievements

### Innovation Challenge Halifax 2019, Sept-2019

Finalist- Ocean Category  
Saint Mary's University

### David Sobey's Annual Retail Hackathon, April-2019

Saint Mary's University (Runner up)

### NS Open Data Contest, March 2019

Dalhousie University (Second Position)