PIZON SHETU

**+1-917-340-3347** [**pizon.skhan@gmail.com**](mailto:pizon.skhan@gmail.com)[**www.linkedin.com/in/pizon-shetu/**](http://www.linkedin.com/in/pizon-shetu/)[**www.github.com/izinex/**](http://www.github.com/izinex/)

# PROFESSIONAL SUMMARY

Dedicated and motivated learner with experience in data analysis, I can extract meaningful insights and build intuitive models to help with decision making, allowing businesses to grow with their objectives in mind using data-driven solutions.

# EDUCATION

## Springboard Data Science Bootcamp

**Online 2021**

Performed and implemented full Python Data Science Stack, Data Wrangling, Statistical Inference, Supervised and Unsupervised Machine Learning, Deep Learning, SQL, A/B Testing, etc.

## Queens College: BA in Computer Science and Applied Math

**NY, Queens 2016 – 2020**

Relevant Coursework: Statistic, Bayesian Modeling, Linear Algebra, Linear Programming, Machine Learning in R.

# TECH STACK

**Python:** Pandas, NumPy, OOP, SciPy, Statsmodel, Tensorflow, Keras, Flask, requests, PyTorch

**Machine Learning:** Neural Networks, Regression, Decision Trees, RandomForest Scikit-Learn, OpenCV, R, Databricks, Classification, RandomForest, NLP, Git

**SQL:** NoSQL, sql-lite, Microsoft SQL Server, DBMS, ETL, Data Analysis, DML

**Apache:** Hadoop, PySpark, MongoDB, MapReduce, HDFS, Hbase, Data mining, Oozie, Mahout

**Visualization:** Tableu, PowerBI, Seaborn, Matplotlib

# RELEVANT EXPERIENCE

### Designed and built Convolutional Neural Network for Image Recognition – Classification

### Utilized Keras API to build a complex neural network which can classify 315 different species of birds.

### Leveraged Transfer-Learning with VGG16, it was able to classify with a 94% accuracy.

### Hyper-parameter tuning the model through Bayesian Optimization and freezing/unfreezing CNN layers led to a further gain of 4% and result in 98% accuracy in predictions.

**Predictive Pricing Model – Regression Model**

* Created a competitive pricing model by analyzing Big Mountain’s attractions and facilities to their competitors.
* My model was able to increase their revenue by $3.4 million.

**New York Housing Price Prediction - Decision Tree**

* Cleaned and implemented MICE imputations on over 75K invalid and missing data for Zillow’s Housing data
* Provided key insights and underlying distribution into NYC housing market, such as wealth gap, cost of homes, and even migration of native residents
* Build a predictive model using Gradient Boosting algorithms with the XGBoost library

### Database for College Classes – SQL

### Build a relational database for Queens College which schedules and maintains classes for the spring semester.

### All classes were properly allocated so no two classes had conflict when it came to time and location.

### Included all course departments and faculty buildings

# PROFESSIONAL EXPERIENCE

## Junior Data Scientist

### ProMarketingHub 2020 - 2021 Queens, NY

* Managed and stored user data on cloud database, performed daily ETL, data cleaning, and preprocessing.
* Coordinated closely with Sr. & Lead Data Scientists to generate and test hypotheses that align with product engagement.
* Assisted Senior team on building website layout based on customer needs using data-driven methods.
* Collaborated and worked on user login authentication using Flask API implementation.
* Defined real-time customer data needs, evaluated data quality, and determined suitability for use.

## Data Analyst

### Centerplate 2016 - 2020 Elmont, NY

* Performed data entries, sorting and analysis on over 10K+ client data to assist with customer engagement.
* Increased customer orders by 13% by incentivizing coupons and combo deals
* Analyzed food data to find most popular orders to restructure pricing and maximize company profits.
* Created and presented data visualization of customer habits and other findings to help further growth.