

# Oyesh Mann Singh

LinkedIn: <http://www.linkedin.com/in/oyeshsingh>

Phone: 4096005899 | Email: [oyeshsin@hotmail.com](mailto:oyeshsin@hotmail.com) | Website: <https://oya163.github.io>

## Education

University of Maryland Baltimore County, MD **08/2017-12/2020**  
Masters of Science in Computer Science GPA-3.756/4

Kathmandu Engineering College, Nepal **05/2007-12/2011**  
Bachelors of Science in Computer Science GPA-70.11/100

## Technical Skills

- **Programming Languages** - Python, R, Java, C++, HTML, CSS, Javascript, Bash
- **Database** - MySQL, PostgreSQL, NoSQL (MongoDB)
- **Others** - Linux, PyTorch, Github, AWS services, Numpy, NLTK, Pandas, Sklearn, Flask, REST API, Docker, Apache Airflow, CI/CD pipeline (Github Actions, Travis CI), Redis, Kubernetes

## Professional Experience

### Research Intern

NAAMII, Kathmandu, Nepal **06/2020-08/2020**

- Developed a non-parametric Bayesian Model for unsupervised morphological segmentation task. Deployed this model as a web app on AWS Elastic Beanstalk following CI/CD pipeline using Github Actions
- Github link - <https://github.com/oya163/gibbs>

### Research Intern

HLTCOE, Baltimore, MD **06/2019-08/2019**

- Increased the entity-tagged dataset volume by 500% by introducing more Wikidata classes and IDF scores in existing Wikipedia framework especially for Chinese and Russian languages
- Improved the NER F1 score by 4% for Chinese language by pre-training BERT with newly obtained dataset compared to the BERT-base

### Data Intern

Relus Cloud Technologies, Atlanta, GA **06/2018-08/2018**

- Successfully completed two major Data Ingestion Pipeline projects on AWS platform to extract millions of real-time records from wearable devices, transform using AWS Lambda scripts and loading into AWS DynamoDB

### Teaching Assistant

UMBC, Baltimore, MD **08/2017-12/2020**

- Assisted 30-40 students in undergraduate courses (Data Structure, Computer Architecture and Data Science) and graduate course (Advance Operating Systems) with excellent feedback from students and professors

### Software Developer

Lamar University, Beaumont, TX **03/2015-05/2017**

- Enhanced Chemical Analyzer Barometer (CAB) software using auto-thresholding technique of Otsu's method to properly detect the edges of water droplets
- Added GUI features in Excel Merge Utility software for more convenience assisting in further data analysis

### Associate Software Developer

Software Paradigm Infotech, Kathmandu, Nepal **01/2013-12/2014**

- Successfully deployed two projects in Credit Management and Reservation domain for Macy's beating the deadlines earlier than expected
- Completed one research project to automatically generate SMS notifications based on voltage load in the building and dump voltage-level logs to database using Arduino. This project was one of the highly appreciated project in the company

## Projects

---

- **Nepali Stemmer:** <https://nepali-stemmer.herokuapp.com> **04/2020**  
Created rule-based stemmer in Nepali language and deployed it on Heroku. Created a pip module for community
- **Aspect Based Abusive Sentiment Analysis:** **09/2019**  
<https://github.com/oya163/nepali-sentiment-analysis>
  - Created Aspect Based Abusive Sentiment dataset in Nepali language
  - This dataset will help to identify who is targeting what kind of abusive words towards whom
  - Trained BERT, BiLSTM and CNN models for Aspect Term Extraction and Polarity Classification task
- **Nepali Named Entity Recognition:** <https://nepner.herokuapp.com/> **02/2019**  
Completed first Nepali NER project achieving 86.9 overall F-score using BiLSTM + Grapheme-level CNN architecture. Deployed this NLP on Heroku using Flask framework
- **Data Science 101:** <https://github.com/oya163/DataScience101> **Oct 2018**  
Participated in Kaggle competition to classify movie reviews and quora insincere questions
- **Data Analysis on 311 Dataset:** <https://github.com/oya163/R-project> **02/2018**  
Led a team of four to analyze and build a predictive regression model to estimate the time taken for a neighborhood based on population of various races in Baltimore using 311 dataset of approximately 3 million records
- **Unsupervised Learning using JigsawPuzzle as pretext task:** **Sep 2017**  
<https://github.com/oya163/oyaTorch/tree/master/JigsawPuzzle>  
Developed unsupervised classification algorithm for medical images based on jigsawpuzzle as a pretext task with accuracy just 5% less than that of supervised

## Certifications

---

- **MongoDB Aggregation Framework** **10/2020**
  - Completed M121: The MongoDB Aggregation Framework course.
  - Gained extensive knowledge on building pipelines and using various aggregation techniques in MongoDB
- **AWS Certified Developer Associate** **11/2020**
  - I am currently taking this course in Udemy
  - Gained extensive theoretical knowledge & hands-on experience with AWS Fundamentals (IAM, EC2, ELB, ASG, RDS, Aurora, S3, CLI, SDK, ECS, ECR, Elastic Beanstalk, CICD pipeline)
- **Apache Airflow** **11/2020**
  - I am currently taking this course in Udemy
  - Creating DAG workflows, running Bash, Python, Postgres and Hive operators
  - Hands-on experience with various executors (Sequential, Local and Celery with RabbitMQ)
- **REST APIs with Flask and Python** **11/2020**
  - Connect web applications to databases and servers via REST APIs
  - Create secure and reliable REST APIs which include authentication, logging, caching, and more
  - Handle seamless user authentication REST APIs and JWT blacklisting
- **Dockers and Kubernetes** **12/2020**
  - I am currently taking this course in Udemy
  - Hands-on experience on building Docker based on multiple containers with Nginx/Redis/Nodejs and designing a CI/CD pipeline from scratch with Travis CI to deploy on AWS EB
  - Configuring Kubernetes cluster and deploying it to the cloud