ASHUTOSH TIWARI

+1 (812)6065974 • checkashu@gmail.com • ashutiwa@iu.edu

LinkedIn Github AnalaticsVidhya Website Medium MachineHack

EDUCATION

COURSE	INSTITUTE/ UNIVERSITY	COHORT	PERFORMANCE
B. TECH. (Computer Sc. and Engineering)	National Institute of Technology, Patna	2011-2015	8.32 (CGPA)
M.S. (Computational Data Science)	Indiana University. Bloomington	Fall 2021-2023	

EMPLOYMENT

1. Swiggy, Software Dev Engineer 2 (Jan 2019 to July 2021)

DATA SCIENCE PLATFORM, DATA ACQUISITION PLATFORM

- Part of Data Science Platform and leading Data Acquisition Platform
- Data Acquisition Platform is a tool used to capture, schedule APIs (mostly third-party or other open APIs). The architecture includes a proxy service, request generation module, transformation module, and scraping module.
- Feature store in DSP provides on-demand features to DSP models in sub ms latencies on a very large scale.
- Worked on **Forecasting and Correlation Platform** which is used by teams to forecast concerning time series.

Spark ML Akka Python Scala

2. Flipkart, Software Development Engineer (ML) 1 (Sep 2017 to Jan 2019)

SEARCH RELEVANCE

- Involved in improvements in search intent model(CRF based) currently in production, identifying error classes, coming up with solutions, and fixing them.
- Worked on Fasttext based query store classifier, which predicts the category of a query, a new model using bi-lstms is going to replace this in near future.
- Wrote the first workflow to automate training and auto-deployment of various search models. First was written using Luigi and later migrated to Airflow.
- Was involved with various POCs with Data Scientists for search intent models.
- Wrote a generic framework using Airflow which at runtime creates generic dags for different ML models and orchestrates their training to deployment flow, including data and model validations.
- Used Cascading to extract data from various user events and then converting it to a form that can be used as training data for our models.

Java ML NLP Python HDFS

3. Groupon, Software Development Engineer 1 (Sep 2016 to Sep 2017)

Seattle based worldwide marketplace

- Cyclops is the interface to Customer Reps which they use to resolve customer queries and requests which is live in all countries in which Groupon operates.
- It talks to almost every microservice in Groupon and also exposes APIs to other internal services.

ROR Backend

4. **NetSpeed Systems,** Software Engineer (Sep 2015 to Aug 2016)

An innovative Silicon Valley SOC simulation startup, acquired by Intel

- Being one of a handful of software developers who participated actively in all development projects.
- Implemented many modules single-handedly (Virtual Channel Arbitration, Polarity based Arbitration, Multi-Cast Filtering, Structural Latency Breakdown, etc.).

C++ | Graph Algorithms | Python

CERTIFICATIONS

COURSE	ORGANIZATION	TIME
Udacity Advanced Machine Learning Engineer Nanodegree	**Udacity** (https://graduation.udacity.com/confirm/9R2E6GC)	**Jan 2019**
P.G. Diploma (Deep Learning)	Indian Institute of Science (IISc), Bangalore (https://drive.google.com/file/d/1E5bN6UrOTZwkIYxxYWPHpnn9p3RT62EW/view?usp=sharing)	Jan-May 2018
External Internship	School of AI (https://drive.google.com/file/d/1-DzDRRXKaoIO067lpvEwu_cTzd3u8-uA/view?usp=sharing)	Mar-Sep 2019
Natural Language Processing with Deep Learning in Python	Udemy(https://naspers.udemy.com/certificate/UC-2b018ead-2e37-46e0-b5ef-cfc70b4e74f7/)	Sep 2020

PROJECTS

- **Flipkart Hackday 9 Ekart Winner Hack 2018**: Build a Multi-label model(inspired by Inception v3) to annotate images of lifestyle products. We also used Google OCR API to extract selected text from the tag. The end goal was to find top candidate FSNs. Text from tag was primarily used for features like price and brand. Others more important ones came from annotations(color, type, cloth type, etc). On top of this to search that product we formed a query using this information and predicted using a CRF model trained on clickstream data of Flipkart using features from features generated using Flipkart's catalog. It was so appreciated that it is in the process of going to production(which is the reason, not providing code pointer here). We did use differential learning rates to tune accuracies in the last stages of training to reach a 99.6% validation accuracy.
- **Groupon Geekon 2017**: Deal recommendations to a user based on NSVD, using it as an unsupervised, collaborative filtering algorithm. Language: Python. Packages: Tensorflow, py2neo, pandas, and NumPy. DB used was Neo4j. Dataset used was movielens dataset. Code here: **https://github.com/thunderock/item_recommendation**.
- **College Major Project:** Selection and simulation of continuous dominant set in case of a distributed sensor network (CDS) and recovery from failure of one and multiple Dominant nodes. Language: C++. Code here: **https://github.com/thunderock/CDSproject**.
- College Minor Project: An android game (A variant of Snake and Ladders). Language: Java(Android Framework). Code here: **https://github.com/thunderock/Snake**.
- Worked on AI democratization program in Swiggy (https://bytes.swiggy.com/executive-espresso-with-head-of-engineering-ai-at-swiggy-dale-vaz-cc7290c8cb78) applying advanced NLP skills on automated FOOD tagging in customer chats with service agents.