**ADITYA VIDHATE**

United States • +1-839-500-4342 • [aditya8497@gmail.com](mailto:aditya8497@gmail.com) • [linkedin.com/in/vidhate/](https://www.linkedin.com/in/vidhate/)

Data Scientist with experience in software development, focused on developing high precision-recall, well-designed & efficient end-to-end workflows with data & models from algorithms to production.

**PROFESSIONAL EXPERIENCE**

ConcertAI, Bangalore July 2020 – Present

**Data Scientist: AI Team**

* Doc classification over 11 classes & 10+ subclasses with >90% precision & recall helping categorize clinical documents; 250M+ docs classified, 1000hrs+ employee time saved, flagship data product made possible
* OpenCV based image preprocessing pipelines to improve precision of Tesseract OCR by 5%+
* Developed zero-shot & few-shot multi-task solver to create end-to-end working prototype models that can be applied on any new dataset quickly; 10+ models supported; 3+ NLP tasks enabled
* Chatbot from scratch for office tasks automation with 7+ intents
* Capturing Biomarkers data and associated numerical, categorical and gene results from clinical notes with >70% precision & recall on most frequent 10 and extendible to any new biomarkers.

ConcertAI, Bangalore July 2019 – June 2020

**Associate Software Engineer: Data Products Team**

* Designed & developed software to speed up Tesseract OCR engine by >50x
* Spark & SQL based pipelines to wrangle Big Data, build ETL pipelines and standardize-normalize data fields
* End-to-end ownership of an NLP tool to support 30 concurrent users for flagship data product
* End-to-end ownership of file MIME type inference algorithm & micro-service
* Designed & developed software for a few-shot learning platform. Config based inference triggering, add-remove-use any model, train pre-existing models with prompts, switch on-off horizontal models

**ACCOMPLISHMENT**

ConcertAI, Bangalore, India 2020

**Above & Beyond Award**

Outstanding performance on algorithms developed and going beyond defined responsibilities to ensure success of a crucial company project.

**EDUCATION**

BITS Pilani Hyderabad Campus - 2019 CGPA 9.0

**MSc (Hons.) Physics, BE (Hons.) Computer Science** (Combined GPA out of 10; Graduated with Distinction)

XII HSC (2014): 93.23%

X ICSE (2012): 94.14%

**SKILLS & OTHER**

Python, Bash, Docker, Multiprocessing, Tesseract OCR, OpenCV, MongoDB, Spark, SQL, FastAPI, Flask

EDA: Regex, K-Means, ANOVA, TfiDf & other vectorizers

Machine Learning: Sklearn, XGBoost

Deep Learning: Spacy, Prodigy, CNN, Word-Vectors, Transformers

NLP Tasks: NER, Classification, Search

Explainable Modelling & Visualization: Matplotlib, SHAP