SENTHIL KUMAR

senthilkumar.m1901@gmail.com • +91-984-186-9609 • GitHub • LinkedIn • Blog • Medium • Chennai, India

- A ML Engineer
  - specialized in building production-grade NLP applications utilizing ML/DL and Generative AI
  - capable of co-developing ML applications by following robust coding practices
  - who has extensively used state-of-the-art Transfer learning models
- An aspiring software engineer who strives
  - to co-develop clean, modular, tested software applications
  - to apply software engineering principles in every datascience/ML effort
- Total Exp: 14 years; Data Science & NLP Exp: 10/14 years

WORK EXPERIENCE

Lead ML Engineer, Toyota Connected India   **Jul 2022 – Present**

- Build data pipelines and NLP applications in AWS cloud to aid Connected Car customers by following Agile Scrum methodology
- Published 4 articles in TCIN medium and presented a seminar on Serverless Python App in PyCon2023

Lead Data Scientist, Ford, Analytics Team   **May 2018 – Jun 2022**

- A hands-on **data science developer** who co-developed with NLP experts from the US team
- I contributed in the end-to-end ML application development
  - from data acquisition, cleaning, labeling and preprocessing,
  - to model development, deployment and maintenance

Assistant Manager, LatentView Analytics (LV)   **Apr 2014 – Apr 2018**

- Roles I played: **NLP Data Scientist**, **Project Delivery Manager**
- Responsible for project scoping and accountable for the delivery of Social Media Analytics projects of 8+ members
- *"....Sincere, driven, articulate and utterly committed ... "* - Skip-level Reporting Manager at LV

Lead Analyst, Beroe Inc   **Jul 2010 – Dec 2013**

- Produce Market Research reports on how to procure indirect spend categories
- *" ... well organized, innovative ... and always ready to go the extra mile"* - Client Engagement Manager

---

Career Progression & Performance Awards

- **Toyota**: Promoted from Senior ML Engineer to Lead ML Engineer in 1.5 years of joining Toyota
- **Ford**: Senior Data Scientist to Lead Data Scientist Promotion
  - In Nov'2019, after 1.5 years of joining Ford
- **LatentView**: Senior Analyst to Assistant Manager Promotion
  - In Oct'16, after 2.5 years of joining LatentView
- **Beroe**: Promoted twice in my first company
  - During my 3.5 year stint in Beroe

- **Toyota**:
  - 1/ Employee of the month. Won in Sep'23, 3 months after joining the company
  - 2/ Hackathon Team Winner. Part of team that won TCIN Hackcelerate in Mar'23
- **Ford**: Asia-pacific Recognition Award
  - Won in May '19 for successful spearheading of a project
- **LatentView**: Encore Award
  - Won for company-wide best performance for the Jul-Sep 2016 quarter
- **Beroe**: Knowledge Contributor Awards (Won twice for company-wide best performance in 2013)

EDUCATION

- Bachelors in Engg - Electronics - 8.6 CGPA
  - Madras Institute of Technology, 2006 - 2010
- 12th Grade - 95% | 10th Grade - 92%
  - State Topper in Physical Science paper, 2006 TN Engineering Entrance Exam

TECHNICAL UPSKILLING

Online Courses
- <u>DevOps on AWS Specialization(4 courses)</u>, Coursera-AWS, Oct'24
- <u>Generative AI with LLMs</u>, Coursera-Deeplearning.ai, Nov'23
- <u>DeepLearning Specialization(5 courses)</u>, Coursera-Deeplearning.ai, Dec'18 - May'19
- <u>ML Fundamentals for Structured Data (2 courses)</u>, Kaggle Learn, Jan-Feb'22
- <u>Applied Text Mining (2 courses)</u>, Coursera-MichiganUniv, Jan'18
- <u>GCP Big Data & ML Fundamentals</u>, Coursera-Google, Apr'21
- <u>SQL (GCP BigQuery) Fundamentals</u>, Kaggle Learn, Feb'22

Skills
- **Languages**
  - Python, SQL(basics), Markdown, Bash/Shell Scripting
- **Python Libraries**
  - Pandas, SpaCy, Re (Regular Expressions), Transformers, Sklearn, PyTorch, OpenAI, LangChain
- **Tools**
  - Git,AWS CLI, Docker, Pipenv (& other Python env management tools), AWS Serverless, OpenAI, ChatGPT, Prompt Engineering

KEY PROJECTS

A Platform with a suite of GenAI Products - X (name withheld) - in AWS | 2024
- Led and worked in a team that built a suite of products in a common platform - X - to unlock actionable intelligence from any text data
  - xWorkFlow - A Workflow Automation Tool with a robust YAML-based prompt engineering system for LLM Chaining in workflows
  - xGPT - An Answer Engine for securely interacting with customer data
  - xAgentChat - A Multi Agent Chat Application armed with access to tools like Calendars and Dealer Systems

Serverless LLM-powered Classification Application | 2023
- Leveraged OpenAI paid models (such as GPT-4) to perform hierarchical text classification
- Utilised AWS Lambda functions to throttle the OpenAI API calls at scale

Contact Center Dashboard - A Speech2Text Dashboard Application in AWS | 2022-23
- Built a ML Pipeline comprising Open Source models converting Call Center audio transcripts to usable insights in dashboard.

BERT Fine-tuned Aspect-based Sentiment Analysis Pipeline | 2020-21   **More Details**
- Built a Sequence Classification ML Pipeline which converts customer comments into trackable Aspect and Sentiment pairs

Personally Identifiable Information (PII) Detection using Named Entity Recognition (NER) | 2020-21   **More Details**
- Annonymized PII in text data by building a NER system using **RoBerta Fine-tuned Transformer model**

NLP Semantic Similarity Engine | 2021-22   **More Details**
- Created a "digital thread" - a NLP-based semantic similarity matching engine based on automotive part descriptions - between two automotive domain specific data sources

Learn by Blogging, My Personal Technical Writing Space   **More Details**
- Have been passionately learning and writing about coding, data science and ML since 2020 <u>here</u>. Many of the 30 articles here are detailed long form blog pieces.