Total experience: 12 years
Bangalore, India
+91 (819) 741 1464
tarunrs@gmail.com
https://github.com/tarunrs

Tarun Ronur Sasikumar

Education

2011 – 2013 **Ohio State University**, *Columbus, OH*.
MS in Computer Science and Engineering. GPA: 3.7/4.0

2004 – 2008 **Vishweswaraiah Technological University**, *Belgaum, India*.
BE in Computer Science and Engineering.

Professional Work Experience

2019 – Present **Technical lead**, *Vahan Inc*, Bangalore, India.
Responsible for backend engineering and mentoring of team members.
Highlights:

- Improved database read time by 2.5x on a very large table by using table inheritance and table partitioning
- Implemented a easily scalable system for websockets increasing throughput by 2x
- Made the internal jobs system more robust by using Apache Airflow reducing failure rate from 20% to 2%.
- Dockerized all products and improved the process of deployment with automated tests.
- Implemented HA on all products to achieve an uptime of 99.5%
- Architected and implemented the backend for the Mitra product
- Build a rule based cohort generation system with custom query generation
- Performed database and query optimizations to reduce query time by orders of magnitude

2016 – 2019 **Data Scientist**, *Crediwatch*, Bangalore, India.
Led the Data Science effort. Responsible for the entire data science pipeline, including system architecture and mentoring of team members.
Highlights:

- Implemented a distributed crawling infrastructure capable of crawling and indexing 100+ pages/minute per bot (using scrapy, frontera, Elasticsearch, Kafka and HBase).
- Designed and developed a real-time news search system with sentiment analysis, named entity extraction and clustering of news stories.
- Built a system to detect and generate Breaking News.
- Improved the precision of a date extractor library by 20%.
- Built multiple classification models to tag news articles with metadata.
- Invented a new pre-clustering algorithm based on Canopy clustering algorithm that also optimizes on the density of the clusters.
- Built a Lead Generation system using machine learning(kNN) and graph mining(Pagerank) techniques with a potential to improve conversion rates by 2-7x
- Built an interactive visualization framework for text such as Adverse News articles similar to quid.com(sigma.js, Elasticsearch)
- Bootstrapped a dashboard framework for analytics team (crontab / python / elasticsearch / HTML / Altair/ Vega)

2013 – 2015 **Software Engineer**, *Epic Systems Corp.*, Verona, Wisconsin.
Responsible for design and development of frontend and backend modules used to document Oncology related information in the EMR. Instrumental in adding support to interface with third-party Radiation Oncology systems.

2008 – 2011 **Software Engineer**, *Atlantis Computing*, Bangalore, India.
Responsible for design, development and maintenance of user profile and application virtualization components in the Atlantis ILIO product stack. Built expertise in the core filesystem level and driver development of both Windows and Linux platforms.

Research Projects

tCTR - Network attention predictor for tweets, Prof. Srinivasan Parthasarathy.
Studied the effect of language syntax and sentiments in sparse data like tweets to predict network attention. Showed that using these lower dimensional language features contributed to network attention more than word based features or novelty of the subject.

Realtime Analysis of Tweets, Prof. Arnab Nandi.
Implemented a system for real-time analysis of sentiments and surfacing of contextual tweets related to political events like the Presidential debates. The supervised model achieves a precision of 85% with a response time of less than 200ms on the 1% Twitter firehose.

Link Prediction - KDD Cup 2012, Prof. Srinivasan Parthasarathy.
Part of a team of 3 that took part in the KDD Cup link prediction challenge in social networks. Wrote multiple scripts to extract semantic and network features from the social graph which helped achieve a weighted average precision of 81%.

pSnipSuggest, Prof. Arnab Nandi.
Implemented a system that provides on-the-go, context-aware assistance in the SQL composition process, based on SnipSuggest by Khoussainova et al. Increased the average precision by 21% and the response time by 30% compared to the original implementation. Written in Python

Personal Projects

jamMm.in.
Co-founded jamMm.in, an online music collaboration application. Served as the User Experience developer. Wrote the framework for mixing multiple tracks as well as the library to generate waveform images for the tracks. Written in Python and Ruby.

These projects are listed with source code at: https://github.com/tarunrs

Technical Skills

Languages	Experience: Python, Javascript, C Intermediate: HTML/CSS, Java, C++ Familiar with distributed/parallel processing systems like Spark, MapReduce/Hadoop ecosystem
Datastores	Postgres, MongoDb, DynamoDb, Elasticsearch, HBase
Datastores	Flask, Node, ExpressJS, NestJs, Django
Message Queues	RabbitMQ, Kafka, Redis