Sriharsha Hatwar

⚲ *Shimoga, India*   ▢ *+91 7019943554*

✉ sriharsha02hatwar@gmail.com   ⌂ Personal Website : sriharsha-hatwar.github.io   ⓖ Sriharsha-hatwar

Publications

- Prajwal Chandrashekaraiah, Pranav kashyap, **Sriharsha Hatwar**, Srinivasa Murthy.
  Lead Artist Identification from Music Videos using Auto-encoders [LINK]
  Symposium on Machine Learning and Metaheuristics Algorithms, and Applications (SoMMA 2019)

News

- (*Sep, 2021*) Our initial idea on **Data augmentation for modality agnostic out of distribution generalization** has been awarded with 500$ credit by **ICLR 2022 DEI Program**.
- (*Aug, 2021*) My task on **Figure of speech detection** has been accepted at GoogleAI's **Big Bench**.
- (*Aug, 2021*) Came in **top 13%** in the Competition organized by Kaggle - **CommonLit Readability Prize**
- (Jan 2021) My notes on different research papers are up and visible at my blog [LINK]
- (*Dec, 2020*) Xperi Hackathon 2020 - Popular choice award - Team **Xvision** (Prize money 25,000 INR)
- (*Jun, 2020*) Attained **bronze medal** for **top 8%** private leader board in the Kaggle competition - **Tweet Sentiment Extraction** - (176th /2227 teams) [KAGGLE-PROFILE-LINK]

Skills and Technologies

- **Programming Languages:** Python, C++, Haxe, C, Java, Javascript
- **Databses & Tech Stack:** MySQL, PostgreS, Redis, PyTorch, HuggingFace, MongoDB, Scikit-learn, Flask, Splunk

Professional experience

Xperi (Formerly Tivo)

Software Engineer - II   *July 2019 - Present*

- Working on UI Enhancment, features and bug fixes in the core UI/UX team that drives features for latest Tivo products. [Linux STBs and soon to launch streamers]
- Implemented Retry mechanism for IP Linear playback for streamers and built Splunk dashboard to monitor the same.
- Redesigned and implemented DeepLinking to Third party apps using Android Intent API.
- Involved in implementation of Subscription management for operators.
- Redesigned TvBeacon Service for in-home device communication and reduced the round trip between mini and host STB for the latest Series 7 Set-top-boxes impacting 10,000+ in-home devices.

Software Engineer - I   *July 2018 - July 2019*

- Exposed to haxe cross-compiler toolkit, allowing one to target both Linux STBs (C++) and modern android based (Java) app for streamers.
- Redesigned UI Progress bar widget across screens, worked on automation tool for sending release notes, perform cherry pick using a webbased interface.
- Implemented the HPK logging for the Bluetooth based remote for enhancing debugging.
- **Languages and Tech Stack** : *(Haxe, C++, C, Python, Java - Android, JS), (Postgres, splunk, Flask, Sqlalchemy, bootstrap)*

Veveo (Subsidiary of Tivo)

Software Engineering Intern   *Jan 2018 - Jun 2018*

- Worked on extracting metadata information from category section in wiki articles and enhanced metadata on media entities by around 3% overall. Also created webpage for comparing metadata results before and after KG creation.
- Implemented a rule based engine to mine information. Later, parallelized the creation a Knowledge graph(KG) by utilizing map-reduce pipelines.
- Improved the KG by utilizing smart tags for each entity by exploiting parent-child relationship between category nodes, created the complete KG by following two pass mechanism.
- **Languages and Tech Stack** : *Python - Spacy, NLTK, Scipy, numpy; HTML, Bootstrap, JS.*

Volunteering

MLCollective (Formerly DLCollective)

Research member   *June 2021 - Present*

- A non-profit ML research organization, where I am working on two collaborative projects.
- Working on researching robust methods for Out of distribution generalization. Additionally, working on blog post submission for ICLR 2022 in collaboration with the Natural-language-processing-subgroup.
- Assisted Rosanne in solving a bug in mlcollective website, later provided a proper fix. [LINK].

In progress Projects

This indicates the work towards publication that is in progress as a member of MLCollective.

Data augmentation for modality agnostic out of distribution generalization   Aug 2021

- This research started as an RFP (Request for plot), an the initiative by DEI at ICLR 2022. I am collaborating with a PhD student. We are working on finding the efficacy of AugMix during data distribution shifts irrespective of the modality. Currently working on WILDS and DomainBed datasets for OOD environments.
- **Tags : (OOD, Robustness in Deep Learning models), (Pytorch, HF transformers, nlpaug, PIL)**

Blog post on "CoCon: A Self-Supervised Approach for Controlled Text Generation"   Nov 2021

- Working on creating a blog post for previous ICLR accepted publication "CoCon: A Self-Supervised Approach for Controlled Text Generation" as a part of **blog posts as conference**. Need to undergo double blind peer review.
- **Tags : (Conditional Language modelling)**

Projects

This section contains selected projects and its brief description. Detailed information at [LINK].

Tweet Sentiment Extraction   Mar - June 2020

- This research competition involved extracting sentiment causing **span of text** from a tweet given text and sentiment of the tweet. Experimented with several transformers and RNN architectures with different pretraining methodologies, added 1-D conv layer on top of transformers, resulted in top 8% and bronze medal in kaggle. **Jaccard Score : 0.7175.**
- **Tags : NLP, Transformers, RNNs, transfer-learning, QA Task, Pytorch, HF transformers**

Clothing Fashion Extraction from Video Songs   Jan - May 2018

- This project involved extracting clothing fashion items from lead artist. Involved Identifying frames containing people, creating a face-body mapping, Identifying lead artist using auto-encoders (Acc : $\sim 70\%$). Finally used trivial Image processing methods to extract clothing using like ; minimum spanning tree segmentation, background elimination.
- **Tags : Computer Vision, Deep learning, Auto-Encoders, Clustering, segmentation, Python-Keras**

FLYLY - A Web based chat application   Mar - May 2017

- Implemented a chat based application similar to slack. Utilized socket.io for realtime push of messages from Node.JS backend, used a MySQL DB for storing and retrieving chat history.
- **Tags: Web-based application, NodeJS, MySQL, socket.io, Bootstrap**.

Finding Semantic similarity between sentences   Mar - Apr 2017

- This is an implementation of a research paper where it computes similarity between short sentences using a semantic knowledge base and also considers word order information implied in the sentences. [PAPER-LINK]
- **Tags : Semantic-knowledge-base, NLTK, Python**

Education

PES University (Formerly PESIT)   July 2014 - Aug 2018

B.Tech in Computer Science (Specialization in Data science)   *Bangalore, India*

Honors: First Class with Distinction (**CGPA : 8.71**)

Major courses undertaken: Data Structures, Operating Systems, Analysis of Algorithms, Algorithms for Information Retrieval, Social Network Analysis, Natural Language processing, Digital Image processing.

Sri Aurobindo Pre University   June 2012 - Apr 2014

*Pre-University Education, Class **12th (581/600 - 96.84%)***   *Shimoga, India*

Sandeepani High School   June 2009 - Mar 2012

*Secondary Education, Class **10th (609/625 - 97.44%)***   *Shimoga, India*

Other Achievements

- **Distinction Award** - Have been awarded distinction in all the eight semesters during my undergrad.
- **522nd rank** (out of 150,000 people) in Karnataka-CET examinations which secured me a spot in PES University.

Certifications

- Deeplearning.ai courses : ***Neural networks and Deep learning***, ***Structuring Machine learning projects***, ***Improving Deep Neural network hyperparamter tuning, Regularization and Optimization***, ***Convolutional Neural Networks***, *CS224n: Natural Language Processing with Deep Learning*
- Certificates of the above courses - [Certificate - 1], [Certificate - 2], [Certificate - 3], [Certificate - 4]