

SHERLY

DATA SCIENTIST

Fields of Interest

- Computer Vision
- Graphs
- Large Scale Distributed Systems

Contact Details

Email:

sher.t92@gmail.com

Phone:

(33) 766106924

LinkedIn:

www.linkedin.com/in/sherlysherly/

Address:

Teknikringen 68A, 31405, Stockholm
11428, Sweden

Nationality/Residency:


Indonesian

Singapore Permanent Resident

Swedish Residence Permit & ID

Links:

 @_sherx

 @xsher

Publications

A Deep Architecture for Multimodal Summarization of Soccer Games, MMSports '19 Proceedings
Proceedings of the 2nd International Workshop on Multimedia Content Analysis in Sports

Languages

English: Written, Spoken (Native)

Chinese: Written, Spoken (Native)

Bahasa Indonesia: Written, Spoken

French: Written, Spoken (Intermediate)

Skills

Python, Scala, C (Elementary), SQL, R, Machine Learning, Deep Learning, Time Series, Natural Language Processing, Computer Vision, Linux, Git, Mac, Docker, Spark, Hadoop, Keras, Tensorflow

Education & Training

KTH Royal Institute of Technology (2019 - Ongoing)

Masters Degree in ICT Innovation - Data Science

Specialisation in Large Scale Distributed Systems.

- EIT Digital - Excellence Scholarship
- Currently conducting a research project on Vehicle Re-Identification with the new dataset published in CVPR 2019.
- Technical courses: Data-Intensive Computing (A), Data Mining and Scalable Machine Learning and Deep Learning

Polytech Nice Sophia Antipolis (2018 - 2019)

Masters Degree in ICT Innovation - Data Science

Grade: 16/20. Mention Très Bien, Highest Honours.

- EIT Digital - Excellence Scholarship
- Published a paper in MMSports'19 Workshop.
- Technical courses: Mathematical Modelling, Big Data Technologies, Data Mining for Networks, Computational Linguistics, Computer Vision, Data Valorization, Time Series, Deep Learning.

National University of Singapore (NUS) (2011 - 2015)

Bachelor in Science, Major in Statistics

- Participated in NUS Overseas College (NOC) - local entrepreneurship programme with overseas immersion at Philadelphia and New York.
- Student Exchange Programme at Université Paris Sud XI (Orsay) in the programme Masters of Mathematical Engineering.

Work Background

Research Intern

Team SPARKS, i3s | Nov 2018 - Jul 2019

- Re-implemented and researched on application of Multiple Instance Neural Network (MINN) on soccer videos for action detection. Project results in a publication in MMSports'19 Workshop.
- Developed Multi-task Learning model with both Regression and Classification tasks for automatic summarisation of soccer videos.

Junior Data Scientist

Go-Jek Singapore | Oct 2017 - Aug 2018

- Improved Query-Ranking model for Food search engine with Natural Language Processing (NLP) techniques.
- Developed Food recommendation system with collaborative filtering, content-based and a hybrid of the two models.
- Customer segmentation and new users insights analysis to drive user adoption rate.
- Built ETL lifecycle, model serving app and manage deployments.

Work Background

Associate Data Scientist

Zendesk Singapore | Apr 2017 - Oct 2017

- Zendesk AnswerBot research and developments to support multiple languages. Deployed to Production.
- Proofs-of-Concept (POCs) of potential of Data Science based features in Zendesk Chat product.
- Built POCs for real time chat sentiment analysis and chat tagging.
- Summarization and topic extraction from Chat dialogues.
- Research and development with the team in Melbourne.

Software Engineer

Zendesk Singapore | Aug 2014 - Mar 2017

- Developed product features: Chat Rating, Operating Hours, and Premium Analytics reports.
- Managed REST APIs for both private and public API.
- Enhanced network protocols to support internal systems.
- Architectural improvements to the servers to support greater loads.
- Spearheaded several architectural and feature enhancements.
- Integration with main Zendesk products: cross-team and cross-country development with teams in San Francisco.

Research and Technical Projects

Robustness of Vehicle Re-Identification algorithms on different datasets

Ongoing

- Analyze the robustness of the existing methods developed for VeRI dataset on unconstrained VeRI-Wild dataset.
- Analyze the underlying conditions within the two datasets that could have contributed to the robustness of the methods by extracting color, texture and semantic features.

A Deep Architecture for Multimodal Summarization of Soccer Games

Nov'18 - Feb'19

- A novel approach for automatic generation of summaries from soccer videos based on multimodal features, including audio energy, event features from sports analytics and visual information from video frames.
- Incorporation of Multiple Instance Learning to alleviate errors arising from the classification of highly homogeneous video data

Image Retrieval Engine

Jan'19 - Feb'19

- Bag-of-Words model and R-MAC (Tolias, et al) to represent images from TRAIN VOC2007 dataset for similar image retrieval.

Re-implementation of Profiling the End Host by Karagiannis et al.

Jan'19 - Feb'19

- Anomaly detection in IP traffic using graphlets in accordance with the paper to perform Support Vector Machine.

Question Similarity in Community Question-Answering, SemEval-2017

Dec'18

- Question relevance classification for community question-answering forums.
- A study the impact of syntactic, semantic and thread-based features on similar questions detection.

Additional Experiences and Skills

Hackathons

- STHLM Tech Fest hackathon - Microsoft Track, Sep 2019,
- Huawei Sweden AI hackathon, 8th-9th Nov 2019.
- Junction 2019, Helsinki, 15th-18th Nov 2019.

Skills

- Violin-enthusiast, outdoor activities (hiking, climbing, etc), dancing.