Sami Alabed

https://samialabed.com linkedin.com/in/samialabed github.com/samialabed contact@samialabed.com

Profile

I am a second-year PhD student at Cambridge University interested in the intersection of Machine Learning and Computer Systems design. My PhD topic is designing computer systems using Bayesian Networks as building blocks. These probabilistic models enable fast and robust auto-tuning with Bayesian Optimization and recovery from failure with counterfactual inference.

Education

Oct 2019 – May 2023 — **University of Cambridge**, *PhD. Computer Science* — **Cambridge, UK**
- Research topic on the optimization of Computer Systems with Bayesian optimization.
- Awarded funding from *The Alan Turing Institute*, the national institute for data science and AI.
- Awarded an *Honorary Cambridge Trust* scholarship.

Oct 2018 – Jun 2019 — **University of Cambridge**, *MPhil. Advanced Computer Science* — **Cambridge, UK**
- Awarded *Cambridge Trust & Students of Cambridge* scholarships.
- Achieved a *Distinction* (80%) in a research master on optimizing large-scale systems using machine learning.

Sep 2014 – Jun 2018 — **The University of Manchester**, *BSc. Computer Science with Industrial Experience* — **Manchester, UK**
- Awarded *Head of School Award* for an outstanding contribution to the life of the School of Computer Science.
- Consistently achieved *First Class* (80%) results in every year.
- Elected as the second-year students' representative.

Experience

Jul 2019 – Oct 2019 — **Twitter,** *Software Engineer Intern - Events Quality* — **London, UK**
- Worked on providing a structure to the event page by analysing and classifying the perspective of tweets.
- Ran several inter-team workshops to onboard new members and explain new tech to previous members.

Real-Time Clustering Algorithm | Scala

Jun 2018 – Sep 2018 — **Twitter,** *Software Engineer Intern - Live Video* — **London, UK**
- Reduced abuse in broadcasts chats by implementing a machine learning-based model to classify messages.
- Visualised the project's impact by designing MapReduce jobs to collect and analyse data from multiple sources.

Scala | Go | SQL | AWS | gCloud

Jun 2017 – Sep 2017 — **Google,** *Site Reliability Engineer Intern - Google Cloud* — **London, UK**
- Automated the detection of users anti-pattern usage of Cloud DataStore by designing a batch job running over petabytes of data.
- Improved site reliability engineers response time to escalations by producing notebooks to visualise anti-pattern usages.

Python | Go | SQL | gCloud | gRPC | SRE

Jul 2016 – Jun 2017 — **Amazon,** *Software Development Engineer - Prime Video* — **London, UK**
- Reduced costs incurred on our team by £2M through optimising our cluster's configurations and host types.
- Developed a low-latency tier-one micro-service responsible for PIN authentication workflow.
- Provided the BI team with real-time analysis by using Kinesis stream into our data warehouse.

AWS | Python | Java | Spring | DevOps

Sep 2015 – Jun 2018 — **The University of Manchester,** *Student Assistant - School of Computer Science* — **Manchester, UK**
- Developed outreach activities using Drones and Oculus to teach programming in schools.
- Integrated the school's labs into AWS and Azure and produced lab manuals.

Azure | NodeJs | OculusRift

Publications

2021 — *Workshop* — **High-Dimensional Bayesian Optimization with Multi-Task Learning for RocksDB.** — **EuroMLSys21**
Sami Alabed, Eiko Yoneki
- Optimized RocksDB IO throughput by x1.3 in ten iterations beating the state-of-the-art by fifty iterations.
- Contributed a method that injects expert knowledge in a multi-task Gaussian Process.

EuroSys | ISBN : 9781450382984 | DOI:10.1145/3437984.3458841

2019 — *Preprint* — **RLCache : Automated Cache Management Using Reinforcement Learning.** — **arXiv**
Sami Alabed
- Improved cache improved cache management decisions by 20% using a multi-agents RL that perfoms eviction, TTL estimation, and caching decisions.

arXiv:1909.13839

Selected Projects

Open Source Projects

Multi-Objective Bayesian optimization in Garbage Collector
EuroSys2020 DW

Improved the read and write latency of Cassandra by 30% compared to baselines by tuning its garbage collector parameters. Constructed the Gaussian Process using ICM kernel to model the multi-task nature of the problem.

PyTorch | Bayesian Optimization | Python

Reinforcement Learning Cache Manager
Master's Thesis, 2019

Optimised cache utilisation by 20% by building a multi-task reinforcement learning powered cache manager to estimate time-to-live and make caching & eviction decision keys. Built a framework to model cache management as a reinforcement learning problem, the framework supported multi-agent and single agent implementations.

GitHub | Tensorflow | Reinforcement Learning | Python

Method Name Prediction
Paper open source implementation, 2019

Implemented a convolutional attention network to summarise Java source code and predict suitable method name based on a published research paper.

GitHub | Tensorflow | Machine Learning | Keras | Python | NLP | CNN | Attention

EdgeComputing : IoT on The Edge
Bachelor's Thesis 2017

Reduced the IoT devices dependency on the cloud and improved their performance by building elastic and reliable edge network by using Docker containers.

Docker | Edge Computing | Python | FaaS | IoT

KalahAi : Boardgame bot
Paper open source implementation, 2017

Developed a self-learning bot to play Mancala inspired by AlphaGo's paper.
Awarded highest score in term of speed and win rate against 13 other bots.

GitHub | Tensorflow | Reinforcement Learning | AlphaGo | Python

Hackathons Projects

Genie : The smart chatbot for everyday services.
Hackathon : HackLondon, 2016

A chatbot written in NodeJs integrated with various online services to ease the process of purchasing online.
Awards : JP Morgan's Hack Choice. NodeJs | Natural Language Processing

MedReader
AMEE Hackathon, 2015

Speed reading app using Java to ease the consumption of medical papers.
Awards : First Place and the Best Use of Elsevier Data. Java | ElasticSearch

Currency&Stock Offline Trader
Warwick University Hackathon, 2014

Financial querying system to enable currency and stock exchange through SMS built using Java.
Awards : Best use of Bloomberg API. Java | SMS

Leadership

Sep 2017 – Jun 2018 | **Peer Assisted Study Session**, *PASS2 Coordinator* | Manchester, UK
- Awarded *Outstanding Contribution to Peer Support* award.
- Improved the employability of students by running algorithms, industry talks, and CV workshops.

Sep 2015 – Jun 2016 | **Microsoft - DevEvangelist**, *Student Partner* | Manchester, UK
- Ranked in the top five Microsoft student partners in impact on Software Development, and Leadership.
- Increased the adaptability of Microsoft's development tools and cloud services by hosting training workshops.

Sep 2015 – Jun 2016 | **HackSoc MCR**, *Co-Founder* | Manchester, UK
- Ran HackSoc focusing on hackathons and tech workshops and won the EU student hackathon league twice.
- Managed a team of 30 students to raise sponsorship for a total of £65,000 and co-organised seven hackathons.

Teaching

Research projects	Model-based Reinforcement Learning in Optimizing Deep Learning Graphs (2019-20). Auto-tuning Spark with Directed Acylic Graph Models in Bayesian Optimization (2019-20).
Developing Labs	Large-Scale Data Processing and Optimization, Data Science.
Courses Supervision	Concurrent and Distributed systems, Data Science, Machine Learning and Bayesian Inference.
Tutorials	Large-Scale Data Processing and Optimization.