

Sami ALABED

<https://samialabed.com> [in linkedin.com/in/samialabed](https://www.linkedin.com/in/samialabed) github.com/samialabed [@ contact@samialabed.com](mailto:contact@samialabed.com)

PROFILE

I am a second-year PhD student at University of Cambridge 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 failures using Causal Contextual Bandit.



EDUCATION

Oct 2019 May 2023	University of Cambridge, PhD. Computer Science Cambridge, UK <ul style="list-style-type: none">> Research topic on the optimization of Computer Systems with Bayesian optimization.> Awarded funding from <i>The Alan Turing Institute</i>, the national institute for data science and AI.> Awarded an <i>Honorary Cambridge Trust</i> scholarship.
Oct 2018 Jun 2019	University of Cambridge, MPhil. Advanced Computer Science Cambridge, UK <ul style="list-style-type: none">> Awarded <i>Cambridge Trust & Students of Cambridge</i> scholarships.> Achieved a <i>Distinction</i> (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 <ul style="list-style-type: none">> Awarded <i>Head of School Award</i> for an outstanding contribution to the life of the School of Computer Science.> Consistently achieved <i>First Class</i> (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 <ul style="list-style-type: none">> 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. <div style="display: flex; gap: 5px;">Real-Time Clustering Algorithm Scala</div>
Jun 2018 Sep 2018	Twitter, Software Engineer Intern - Live Video London, UK <ul style="list-style-type: none">> 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. <div style="display: flex; gap: 5px;">Scala Go SQL AWS gCloud</div>
Jun 2017 Sep 2017	Google, Site Reliability Engineer Intern - Google Cloud London, UK <ul style="list-style-type: none">> 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. <div style="display: flex; gap: 5px;">Python Go SQL gCloud gRPC SRE</div>
Jul 2016 Jun 2017	Amazon, Software Development Engineer - Prime Video London, UK <ul style="list-style-type: none">> 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. <div style="display: flex; gap: 5px;">AWS Python Java Spring DevOps</div>
Sep 2015 Jun 2018	The University of Manchester, Student Assistant - School of Computer Science Manchester, UK <ul style="list-style-type: none">> 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. <div style="display: flex; gap: 5px;">Azure NodeJS OculusRift</div>

PUBLICATIONS

2021  Workshop	High-Dimensional Bayesian Optimization with Multi-Task Learning for RocksDB. EuroMLSys21 <i>Sami Alabed, Eiko Yoneki</i> <ul style="list-style-type: none">> 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. <div style="display: flex; gap: 5px;">EuroSys ISBN : 9781450382984 DOI:10.1145/3437984.3458841</div>
2019  Preprint	RLCache : Automated Cache Management Using Reinforcement Learning. arXiv <i>Sami Alabed</i> <ul style="list-style-type: none">> Improved cache improved cache management decisions by 20% using a multi-agents RL that performs eviction, TTL estimation, and caching decisions. <div style="display: flex; gap: 5px;">arXiv:1909.13839</div>

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](#)

EDGE COMPUTING : 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 > Awarded <i>Outstanding Contribution to Peer Support</i> award. > Improved the employability of students by running algorithms, industry talks, and CV workshops.	Manchester, UK
Sep 2015 Jun 2016	Microsoft - DevEvangelist, Student Partner > 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.	Manchester, UK
Sep 2015 Jun 2016	HackSoc MCR, Co-Founder > 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.	Manchester, UK

TEACHING

Research projects	Model-based Reinforcement Learning in Optimizing Deep Learning Graphs (2019-20). Auto-tuning Spark with Directed Acyclic 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.