

SUCHIR SALHAN

MEng (CANTAB) COMPUTER SCIENCE STUDENT · MACHINE LEARNING · NATURAL LANGUAGE PROCESSING
STARRED FIRST BA (CANTAB)

🌐 <https://www.suchirsalhan.com/> ✉ sas245@cam.ac.uk
📞 07946290505 🌐 www.linkedin.com/in/ssalhan 📄 ResearchGate

I'm Suchir Salhan, a fourth-year MEng Computer Science student at Gonville and Caius College, the University of Cambridge, specialising in Machine Learning and Natural Language Processing. My principal research interests focus on engineering more data-efficient “foundation models” (particularly Large Language Models and Vision-Language Models). I have a broad interest in Machine Learning Systems, the Theory of Deep Learning and Theoretical Linguistics. I have worked with several researchers in the Department of Computer Science & Technology, particularly in the Natural Language and Information Processing Group (NLIP).

EDUCATION

Oct 2020 - Jul. 2024

MEng in Computer Science

GONVILLE & CAIUS COLLEGE, UNIVERSITY OF CAMBRIDGE

- **Part III Computer Science Tripos:** Admitted to the integrated MEng/Part III of the Computer Science Tripos after a highly-competitive interview with Deputy Head of Computer Science, Prof Timothy Jones, and Russell Moore. **Class I with Distinction** aggregate mark in my undergraduate studies at Gonville & Caius College, University of Cambridge.
- High First Class in every Tripos Paper in Finals. 75+ in Linguistic Theory, Syntax/Semantics, Typology, Morphology and Phonological Theory and Computational Linguistics.
- **Modules:** Machine Learning Modules include **Advanced Topics in Machine Learning** and **Machine Learning for Language Processing**. Advanced Modules in Natural Language Processing, including **Natural Language Syntax and Parsing** and **Introduction to Computational Semantics**. Computer Security (Cybercrime).
- **Part III Project: Less is More: Augmenting Small-Scale Language Models with Syntactic Inductive Biases.** I am currently working on a self-proposed project that aims to engineer **small-scale, compute-efficient Transformer-based Language Models**. With Prof Paula Buttery, Dr Andrew Caines, Richard Diehl Martinez & Zebulon Goriely in the Natural Language and Information Processing Group.
- **Part IIB Dissertation Published**, awarded a High First Class (75+).
- **Examiner's Feedback on Dissertation:** “This is an **outstanding dissertation** which presents a new approach to the topic of computational language modelling: incorporating theoretical principles from the work of Theresa Biberauer (from the model of ‘Maximise Minimal Means’) into the framework of modern-day ‘Transformer’ neural networks. The argumentation and discussion are **sophisticated and accomplished**, the point of view is confident. This dissertation could serve as the basis for a future graduate research project which would empirically test the predictions made in this work. The conjunction of linguistic theory and computational modelling is a relatively neglected area which has great potential for exploration, as proposed and demonstrated in this dissertation. There is evidence of extensive background reading and research, supplemented with thorough critical analyses in which the relevant literature is handled with dexterity and acumen. Presentation and referencing are immaculate. The addition of a case study also demonstrates strong practical skills, the ability to handle large datasets, and perform analyses at the intersection of computational method and linguistic theory.”
- I achieved a **first class with distinction mark (83, 98% in practical assessment)** in Part IIA Tripos examination paper covering Natural Language Processing and Computational Linguistics. Examiners' comments stated that my script was “remarkably complete” and can serve as “a basis for a literature review in a piece of published research”.
- **Seminars:** Syntax Lab, Natural Language and Information Processing (NLIP) Seminars, Language Technology Lab (LTL) Seminars. Co-organised a student seminar and reading group for Phonological Theory, with Prof Bert Vaux. Audited Graduate Seminar in Optimality Theory.

 PUBLICATIONS

CONFERENCE AND JOURNAL PUBLICATIONS

1. †**Salhan, S.A.** (2023) **On the Potential for Maximising Minimal Means in Transformer Language Models: A Dynamical Systems Perspective.** In Cambridge Occasional Papers in Linguistics. Volume 15, Article 3: pp. 55–110. ISSN: 2050-5949. Available online here.
2. **Salhan, Suchir,** Liu, Fangyu & Collier, Nigel. (2022/forthcoming) **Multimodal Language Modelling across Languages and Cultures: Grounding Strategies for Nouns and Verbs.** (Research Project, Language Technology Lab, Department of Theoretical and Applied Linguistics, University of Cambridge). [partially supported by a research award from Gonville & Caius College, Cambridge]

SEMINAR PRESENTATIONS AND WORKING PAPERS

3. **Salhan, S.A.** (2023) **On the Potential for Maximising Minimal Means in Transformer Language Models.** Talk presented at SyntaxLab (14th February 2023), Department of Theoretical and Applied Linguistics, a departmental seminar organised by Dr Theresa Biberauer.
4. **Salhan, S.A.** (2021). **UROP Project Report 2021: Providing Automatic Feedback on Argumentation Quality to Learners of English** (UROP Research Internship, Natural Language and Information Processing Group, Department of Computer Science and Technology, University of Cambridge). [supervised by Prof Paula Buttery, Dr Andrew Caines, Dr Russell Moore (NLIP Group); Dr Thiemo Wambsganss (University of St Gallen/EPFL) and funded by Cambridge Assessment]

 EXPERIENCE

2022	Language Technology Lab RESEARCHER I undertook an extended research project with Prof Nigel Collier and Fangyu Liu (Google DeepMind) in the Language Technology Lab (University of Cambridge). We worked on a creative probing of state-of-the-art multimodal language models. This research has, in part, been supported by Gonville and Caius College.
2021	Department of Computer Science & Technology RESEARCH ASSISTANT I supported my supervisor Dr Li Nguyen (University of Cambridge) as a research assistant in the drafting of her computational sociolinguistics paper on variation in overt/null pronouns by Vietnamese-English bilingual speakers.
2021	Department of Computer Science & Technology UNDERGRADUATE RESEARCHER I was offered a UROP research project by my Director of Studies Prof Paula Buttery and Dr Andrew Caines. I worked under the supervision of computational linguists from the Automated Language Teaching and Assessment (ALTA) group in the Department of Computer Science and Technology. I was the only first-year student admitted to the two-month UROP programme. I worked on a project in collaboration with Thiemo Wambsganss on developing the back-end machine learning architecture for an application that supports the argumentation skills of English language learners. I trained and evaluated the performance of state-of-the-art transformer language models on downstream argumentation mining tasks, began working on the deployment of the pre-trained model in the application, and submitted the ethics application for the experimental evaluation of the educational outcomes of the application. I presented my work to members of the ALTA group and the project sponsors from Cambridge Assessment who funded my UROP project. I also had the opportunity to attend machine learning classes, and seminars on Dialogue Systems, Ethics in NLP, active learning paradigms and educational technology.

AWARDS & SCHOLARSHIPS

- 2023 GONVILLE & CAIUS COLLEGE
Memorial Prize, one of two academic accolades awarded to Gonville & Caius graduating cohort 2023
- 2023 GONVILLE & CAIUS COLLEGE
Examination Prize
- 2023 GONVILLE & CAIUS COLLEGE
Re-election to Senior Scholarship
- 2022 GONVILLE & CAIUS COLLEGE
Senior Scholarship
- 2022 GONVILLE & CAIUS COLLEGE
Examination Prize for First Class Tripos Performance
- 2022 GONVILLE & CAIUS COLLEGE
Research Award awarded by the Senior Tutor of Gonville and Caius (Dr Andrew Spencer), for my research project with Prof Nigel Collier.
- 2022 COMPUTER SCIENCE DEPT.
 Funding of £175 for **Research Assistant** work
- 2021 COMPUTER SCIENCE DEPT.
 Funding by Cambridge Assessment for a UROP in the Department of Computer Science & Technology, worth £3000. I was the only first-year student in the entire University to participate in the two/three month Natural Language Processing UROP program.

EXPERIENCE

- 2022 Language Technology Lab RESEARCHER
 I undertook an extended research project with Prof Nigel Collier and Fangyu Liu (Google DeepMind) in the Language Technology Lab (University of Cambridge). We worked on a creative probing of state-of-the-art multimodal language models. This research has, in part, been supported by Gonville and Caius College.
- 2021 Department of Computer Science & Technology RESEARCH ASSISTANT
 I supported my supervisor Dr Li Nguyen (University of Cambridge) as a research assistant in the drafting of her computational sociolinguistics paper on variation in overt/null pronouns by Vietnamese-English bilingual speakers.
- 2021 Department of Computer Science & Technology UNDERGRADUATE RESEARCHER
 I was offered a UROP research project by my Director of Studies Prof Paula Buttery and Dr Andrew Caines. I worked under the supervision of computational linguists from the Automated Language Teaching and Assessment (ALTA) group in the Department of Computer Science and Technology. I was the only first-year student admitted to the two-month UROP programme. I worked on a project in collaboration with Thiemo Wambsgans on developing the back-end machine learning architecture for an application that supports the argumentation skills of English language learners. I trained and evaluated the performance of state-of-the-art transformer language models on downstream argumentation mining tasks, began working on the deployment of the pre-trained model in the application, and submitted the ethics application for the experimental evaluation of the educational outcomes of the application. I presented my work to members of the ALTA group and the project sponsors from Cambridge Assessment who funded my UROP project. I also had the opportunity to attend machine learning classes, and seminars on Dialogue Systems, Ethics in NLP, active learning paradigms and educational technology.

 SKILLS

- **Programming Languages:** Python (proficient); basic experience in MATLAB.
- **Machine Learning Frameworks:** NumPy, SciPy, matplotlib, keras, scikit-learn, tensorflow
- **Natural Language Processing Frameworks/Libraries:** NLTK, SpaCy, HuggingFace libraries

 INDUSTRY AND OTHER ACTIVITIES

2023-	Co-Founder and Editor-in-Chief of Per Capita Media Per Capita is a progressive national student publication founded in Cambridge University in collaboration with University of Art's London, covering News, Features and Culture. The publication is supported by Lady Stothard, Dr Ruth Scurr. President of Cambridge University Per Capita Media.	MEDIA & JOURNALISM
2024	Founder and President of Gonville & Caius Media and Journalism Society	MEDIA & JOURNALISM
2023-2024	Editor & Head of Events Editor of Cybersecurity Policy Paper for policy thinktank , The Wilberforce Society. Published in Nov 2023 and presented in St John's College, available online here. Organising Freedom of Press Panel Event in St John's College and a panel event with the Foreign Minister of Sri Lanka.	POLICY
2024	The Cambridge Student Deputy Editor— the student paper of Cambridge University.	MEDIA & JOURNALISM
2023	The Cambridge Student News & Investigations Editor and Interviews writer— the student paper of Cambridge University.	MEDIA & JOURNALISM
2023	BBC The One Show Invited by producers at The One Show to help produce a documentary about higher education post-COVID. Responsible for organising student contributors around the country. Also, invited by producers at BBC Radio 5Live to discuss the Marking and Assessment Boycott.	MEDIA & JOURNALISM
2023	Radio Presenter Qualified Radio Presenter at CamFM.	MEDIA & JOURNALISM
2022-23	Varsity Newspaper Member of the Editorial Team (Investigations Editor, Senior Science Editor, News Correspondent, 'Chief Sub-editor') of Varsity— the student paper of Cambridge University.	MEDIA & JOURNALISM
2022	Gonville & Caius College JCR I was elected Secretary of Caius JCR in 2022-23	COMMUNITY
2021-2024	University of Cambridge CamFM, Caius Boat Club (Novice), Member of Cambridge Law Society (CULS), Cambridge Investment Banking Society (CIBS), Cambridge Union, Sub-editor and Author of Polyglossia Magazine 2020-2021, Member of Gonville and Caius College Access Team (helping in open days, access events).	SOCIETIES