Sheikh Abdur Raheem Ali

http://abdur-raheem.com

Email : ali@abdur-raheem.com

Mobile : +1 (604) 790 5369

EDUCATION

Stanford University — Palo Alto, CA, USA
Advanced Software Systems, Dropped out after one Quarter, Moved to Microsoft Canada — *Dec 2020 – Apr 2021*

American University of Sharjah — Sharjah, UAE
Bachelor of Science in Computer Science; Minor in Applied & Computational Math — *Sep 2017 – Dec 2020*

EXPERIENCE

Independent — Toronto, ON, Canada
Research Engineer — *May 2024 – Present*
- Working on the AI alignment problem.

Microsoft — Vancouver, BC, Canada
Software Engineer II — *Apr 2021 – May 2024*
- Youngest engineer ever hired to my level in Canada. Typically requires a PhD or equivalent experience.
- Co-owned impersonation protection in Defender for Office 365. Worked with the FBI, NYPD, U.S DoD to protect 250m+ people against phishing threats. Cut mailbox intelligence false positive rates by 55%.
- Reduced download time for Substrate (the #2 largest monorepo in MS) from 2-3 hours to $\approx$ 20 minutes. Used by over 10% of all developers at the company and saves at least \$1m per annum.
- Climbed to rank #14/36,210 members on our internal bug reporting platform. My feedback has been used by 86 teams (including Github Copilot, Bing Chat, Azure OpenAI) to fix issues before customer impact.

Software Engineering Intern — *Jul. 2020 – Apr 2021*
- Built priority accounts. Was asked to stay on contract after the end of my internship until I graduated.

SWVL — Swvl Headquarters, Dubai, UAE
Quantitative Finance Intern — *Feb 2020 – Apr 2020*
- Rewrote the system that prices how much a ride-sharing startup unicorn charges for each transaction.

Dubai Electronic Security Center — Cyber Intelligence Lab, University of Dubai, Dubai, UAE
ML Research Intern — *Dec 2019 – Jan 2020*
- Recruited by the Government's Head of Research & Innovation. Did a 5-week project in my first weekend.

PROJECTS

- **Scaling Laws for Refusal Mechanisms in Llama-2-70b**: Accepted poster at (TAIS 2024). Produced as part of my stream in Berkeley AI Safety's Supervised Program in Alignment Research (SPAR) - Spring 2024 Cohort. In this period my students made submissions to NeurIPS, EMNLP, ICML MI Workshop, and are working on a paper for ICLR 2025.
- **Interpretable Forecasting with Transformers**: Eliciting accurate, well-calibrated, and honest predictions from gpt-4-base. Supported by independent grants and mentored by Nuño Sempere. External partnership with SERI MATS (Cyborgism stream) building high-bandwidth interfaces for human-LLM collaboration. More details at manifund page.

AWARDS & SCORES

- **2nd place (USD 35,000) @ HITB+ CyberWeek AI Challenge**: Captain of a team including students from Harvard and UToronto that used genetic algorithms to modify existing malware to defeat virus detection agents.
- **Full Scholarship @ European Summer Program on Rationality**: Invited to the UK to attend seminars on topics related to applied mathematics, AI safety, and cognitive science. Taught by ex-Google employees and MIT graduates.
- **Perfect Score @ ACT**: First attempt. Out of 2.03m students who took the 2017 ACT, about 2000 got a full 36/36.

SKILLS

- Transformer Inference, Torch Streams, Pipelining Systems, Distributed Profiling, Python Dataclasses, Information Security, Content Filtering, Developer Tools, Data Visualization, Performance Optimization