Alex MATHAI
Computer Science Researcher

⚲ alexmathai.online   🎓 Google Scholar   in linkedin.com/in/alex-mathai-403117131
@ alexmathai98@gmail.com   📱 +91 9983560468   📍 Mumbai, Maharashtra

❤ RESEARCH INTERESTS

I am passionate about solving challenging research problems. My areas of interest include **machine learning for structured data like programming languages and graphs**, **knowledge graphs** and **graph neural networks**.

🎓 EDUCATION

Years	Degree	Grade	Rank
'16-20	B.E. in Computer Science, Birla Institute of Technology and Science Pilani	9.63 / 10	6th/123

🏆 AWARDS AND SCHOLARSHIPS

Years	Award
'16-20	**Institute Merit Scholarship** Awarded to top 3% students for Exceptional Academic Performance

📖 RESEARCH PAPERS

MONOLITH TO MICROSERVICES : REPRESENTING APPLICATION SOFTWARE THROUGH HETEROGENEOUS GNNS   IBM RESEARCH

Alex Mathai, Sambaran Bandyopadhyay, Utkarsh Desai, and Srikanth Tamilselvam.
Accepted at IJCAI 2022 : ⇱ Paper Link.

INCREMENTAL ANALYSIS OF LEGACY APPLICATIONS USING KNOWLEDGE GRAPHS FOR APP MODERNIZATION   IBM RESEARCH

Saravanan Krishnan, **Alex Mathai**, Amith Singhee, Atul Kumar, Shivali Agarwal, Keerthi Narayan Raghunath, David Wenk.
Accepted at ACM CODS-COMAD 2022 : ⇱ Paper Link.

RECLIVE : REAL-TIME CLASSIFICATION AND QOE INFERENCE OF LIVE VIDEO STREAMING SERVICES   UNSW SYDNEY

Sharat Chandra Madanapalli, **Alex Mathai**, Hassan Habibi Gharakheili, and Vijay Sivaraman.
Accepted at IEEE IWQOS 2021 : ⇱ Paper Link.

ADVERSARIAL BLACK-BOX ATTACKS ON TEXT CLASSIFIERS USING GENETIC OPTIMIZATION GUIDED BY DNNS   IBM RESEARCH

Alex Mathai, Shreya Khare, Srikanth Tamilselvam, Senthil Mani
Available at Arxiv 2020 : ⇱ Paper Link.

💼 US PATENTS

HETEROGENEOUS GRAPH GENERATION FOR APPLICATION MICROSERVICES, (IBM RESEARCH 2021)   PATENT NO : P202101668US01

A patent for the creation of a heterogeneous graph from an application codebase.
Clustering is then performed on this graph to create multiple smaller clusters which represent microservices.

SYSTEM AND METHOD TO EXPLAIN CANDIDATE MICROSERVICES FROM MONOLITH, (IBM RESEARCH 2021)   PATENT NO : P202104631US01

A patent for the explanation of the microservices generated from a monolith.
Adapted GNN explanability models like **GNN explainer** for unsupervised graph clustering

💼 INDUSTRIAL AND ACADEMIC RESEARCH EXPERIENCE

Current Aug 2020	Research Engineer, IBM Research | Modernizing Monoliths to Microservices , BANGALORE, India › Worked with ⇱ Amith Singhee on the knowledge graph (KG) modelling of large monolith applications, and with ⇱ Srikanth Tamilselvam on KG partitioning to generate candidate microservices. › This project resulted in **two US patent** filings and publications at **IJCAI 2022** and **CODS-COMAD 2022**. › Played a pivotal role in (i) **KG construction** using static analysis, (ii) **KG Inference** using graph traversal algorithms and (iii) **KG Partitioning** using graph neural network based representation learning coupled with unsupervised clustering to generate groups (microservice recommendations). Java | Python | Neo4j | SQL | Pytorch

Dec 2019 **Aug 2019**	University of New South Wales (UNSW) | Time Series Analysis Research, Sydney, Australia › Worked with ⬀ Vijay Sivaraman and ⬀ Hassan Habibi on creating deep learning models that can **classify** and isolate **live streaming** network traffic in **real-time** for 5000 Australian homes. › This work concluded with a publication at **IEEE IWQoS 2021**. › Played a lead role in crafting and **extracting time-series features** at gigabyte scales. These features were resilient to challenging settings like network congestion and varying bandwidth capacities. `Pytorch` `Python` `Numpy` `Matplotlib`

Aug 2019 **May 2019**	IBM Research Labs | Natural Language Research Engineer, Bangalore, India › Generated **adversarial text** using black-box techniques to break NLP models trained for classification. › Detailed the entire approach in an arxiv paper. › Integrated deep learning models with a genetic algorithm for a hybrid AI system. › Implemented LSTM classifiers, sequence to sequence translation models and language models. `Pytorch` `Python` `NLTK` `Numpy` `Matplotlib` `FairSeq` `TorchText` `InferSent` `GloVe`

Aug 2018 **May 2018**	Indian Space Research Organization | Computer Vision Research Engineer, Dehradun, India ⬀ Code ⬀ Description › Worked with ⬀ Ashutosh Kumar Jha and ⬀ Sameer Saran to create an **Image Segmentation API** for high-resolution satellite images. › Implemented **U-Nets** for the segmentation of tarred roads and deep water bodies. › Preprocessed 500 GB of multi-spectral geospatial data for its use in the computer vision model. `Tensorflow` `Python` `Matplotlib` `Shapely` `Gdal` `Osr` `Pandas` `Descartes`

May 2018 **Dec 2017**	Web Intelligence and Social Computing Lab | Customer Review Sarcasm Research, Pilani, India ⬀ Code ⬀ Description › Worked with ⬀ Yashvardhan Sharma to develop a system capable of **detecting sarcasm** in online customer reviews. › Used **convolutions for text** to detect personality traits and the underlying emotion of the customer. › Helped replicate, optimize and improve upon models developed by NTU, Singapore. `Tensorflow` `Python` `Pandas` `Numpy` `Matplotlib` `NLTK` `GloVe`

</> Technical Expertise

Programming	Python, Java
Frameworks	Pytorch, Tensorflow, Flask-restx
Databases	Microsoft SQL Server, Neo4j Graph DB
Operating Systems	Mac OS X, Windows 7/8/10, Linux Redhat, Linux Ubuntu

👥 Leadership Roles

2018-19	**Machine Learning Special Interest Group Coordinator, ACM** BITS Pilani Chapter (India's best ACM chapter 4 years in a row)

👥 Teaching and Mentoring

Jan 2020 **May 2020**	Neural Networks and Fuzzy Logic Teaching Assistant, BITS Pilani, India › I was the head coordinator overlooking 10 other teaching assistants. › I helped create meaningful assignments and stimulating quizzes for over 150 students. `Pytorch` `Numpy` `IPython`
May 2019 **Jan 2019**	Database Systems Teaching Assistant, BITS Pilani, India › Prepared material and conducted biweekly lab sessions on Microsoft SQL Server for 235 students. › Helped in framing questions for evaluation components and in evaluating answer scripts. `SQL`