

Muhammad Yaseen Khan · محمد یاسین خان

Curriculum Vitae

Research & Development Department
Love for Data
Pakistan

Tel: +92-332-3240104
Email: yaseenkhan.yzai@gmail.com
Skype: muhammad_yaseen_khan
URL: Google Scholar · ResearchGate · GitHub

Educational Background

Ph.D. Mohammad Ali Jinnah University, 2018–cont'd. Advisors: Dr. Tafseer Ahmed and Dr. Shaukat Wasi
MS.SE Karachi Institute of Economics & Technology, 2014–2015. Advisor: Dr. Khuram N. Junejo
BS.SE University of Karachi, 2007–2010. Advisors: Dr. Humera Tariq and Dr. Tahseen A. Jilani

Position Held

Feb. 2021–Present Principal R&D Engineer (for NLP) at Love for Data, Karachi, Pakistan.
Jan. 2017–Feb. 2021 . . . Senior Lecturer, Department of Computer Science, Faculty of Computing, Mohammad Ali Jinnah University, Karachi.
Jan 2019–Dec. 2019 . . . AI/Software Developer, AksIQ.
Jan. 2016–Mar. 2017 . . . Data Scientist, Cubix Co.
June 2013–Dec. 2016 . . . Lecturer, College of Computing and Information Sciences, Karachi Institute of Economics and Technology.
Feb. 2012–June. 2013 . . . Software Engineer II, Mazik Global Technologies / Microsoft Inc. / Tyler Technologies Inc.
Jan. 2011–Feb. 2012 . . . Web and iOS Application Developer, Counterpoint/Qurta LLC.
July 2010–Dec. 2010 . . . Web Developer, Revolutionary Technologies.
Jan. 2010–June 2010 . . . Web Developer, Sprintech Solutions.

Teaching

Fall 2020 CS6420 Applied Programming; CS1210 Introduction to Computing
Summer 2020 CS6420 Applied Programming;
Spring 2020 CS6420 Applied Programming; CS3310 Artificial Intelligence
Fall 2019 CS6420 Applied Programming; CS1420 Object-Oriented Programming
Summer 2019 CS1410 Computer Programming; CS3110 Software Engineering
Spring 2019 CS6420 Applied Programming; CS6610 Information Retrieval Techniques (co-taught with Dr. Shaukat Wasi); CS2510 Data Structures and Algorithms;
Fall 2018 CS2510 Data Structures and Algorithms
Spring 2018 CS1410 Computer Programming
Fall 2017 CS1410 Computer Programming
Spring 2017 CS1410 Computer Programming; CS1420 Object-Oriented Programming; CS5339 Introduction to Data Science; CS5510 Advanced Research Method (co-taught with Dr. Shaukat Wasi)
2013–2016 Courses taught to undergrad students at KIET in different semesters: Programming Fundamentals; Object-Oriented Programming; Data Structures and Algorithms; Data Mining; Software Engineering.

Conference and Workshop Committees

1. Organising committee, 1st International Conference on Data Sciences, MAJU, Karachi (2019)
2. Trainer, 1-Day Workshop on Data Science with Python, UBIT, Karachi University (2018)
3. Trainer, Hands-on Programming with Python, MAJU, Karachi (2018)

4. Trainer, Programming with Python, Combat:18, KIET, Karachi (2018)
5. Coach, ICPC 2015 @ NUCES-FAST, Lahore (2015)
6. Faculty head, CodeJam, PAF-KIET, Karachi (2015)
7. Faculty head, Combat: The Computing Battle 2017, KIET, Karachi (2014)
8. Coach, Startup-Weekend, IBA, Karachi (2012)

University Committees

1. Examination paper moderation committee, MAJU, Karachi (2018–present; *chair*, Spring 2020)
2. Committee on examination policy for COVID-19 pandemic, MAJU, Karachi (*chair*, Spring 2020)
3. QEC Self-Assessment Report for undergrad and grad CS and SE programmes, MAJU, Karachi (member, 2019)
4. External course review committee, KIET, Karachi (2019)
5. Graduate thesis evaluation committee, MAJU, Karachi (member, 2018–present)
6. Academic advisor, MAJU, Karachi (2017–present)
7. Final year project evaluation/exhibition committee, MAJU, Karachi (member, 2017–present)
8. Founding faculty sponsor of ACM-MAJU student chapter, MAJU, Karachi (2017–present)
9. Admission interviews committee, MAJU, Karachi (member, 2017–present)
10. Credit transfer committee, MAJU, Karachi (member, 2018–present)
11. Faculty sponsor of ACM-KIET student chapter, KIET, Karachi (2013–2015)
12. Final year project committee, KIET, (member, 2013–2016)

Boards

1. Board of Studies, MAJU, Karachi (member, 2017–2019)

Publications

Items are sorted chronologically in reverse order. Conference papers that later converted into book chapters or journal articles are counted once.

Refereed Journal Articles

1. “Hindustani or Hindi vs. Urdu: A Computational Approach for the Exploration of Similarities Under Phonetic Aspects”. Muhammad Suffian Nizami, Tafseer Ahmed, and [Muhammad Yaseen Khan](#); *International Journal of Advanced Computer Science and Applications*, volume 11, number 11, pages 749–755, 2020. (WoS ESCI indexed). DOI: 10.14569/IJACSA.2020.0111191
2. “Edit Distance-based Search Approach for Retrieving Element-wise Prosody in Hindi/Urdu Poetry”. [Muhammad Yaseen Khan](#), Muhammad Adil Rao, Shuakat Wasi, Twaha Ahmed Minai; *Indian Journal of Science and Technology*, volume 13, number 39, pages 4189–4201, 2020. (MJL indexed). DOI: 10.17485/IJST/v13i39.1489
3. “Exerting 2D-Space of Sentiment Lexicons with Machine Learning Techniques: A Hybrid Approach for Sentiment Analysis”. [Muhammad Yaseen Khan](#) and Khurum Nazir Junejo; *International Journal of Advanced Computer Science and Applications*, volume 11, number 6, pages 599–608, 2020. (WoS ESCI indexed). DOI: 10.14569/IJACSA.2020.0110672.
4. “Sentiment Analysis of Product Reviews in the Absence of Labelled Data using Supervised Learning Approaches”. Warqar Muhammad, Maria Mushtaq, Khuram Nazir Junejo, and [Muhammad Yaseen Khan](#); *Malaysian Journal of Computer Science*, volume 32, number 2, pages 118–132, 2020. (WoS SCI indexed). JCR Impact Factor: 0.7. DOI: 10.22452/mjcs.vol33no2.3
5. “Depression Detection using Naïve Bayes: Text based approach”. Shah Muhammad Emad, Kamran Ahsan, [Muhammad Yaseen Khan](#), Shoaib Ahmed, Rafi Ullah, Kashif Rizwan, and Farhan Shafiq; *Interciencia*, volume 44, number 11, pages 174–182, 2020. (WoS SCIE indexed). JCR Impact Factor: 0.26.
6. “On Extraction of Event Information from Social Text Stream: An Unpretentious NLP Solution”. Kanwal Iqbal, [Muhammad Yaseen Khan](#), Shaukat Wasi, Shumaila Mahboob, and Tafseer Ahmed; *International Journal of Computer Science and Network Security*, volume 19, number 9, pages 121–131, 2019 (WoS ESCI indexed).

7. "A Novel Approach For Ontology-Driven Information Retrieving Chatbot For Fashion Brands". Aisha Nazir, [Muhammad Yaseen Khan](#), Tafseer Ahmed, Syed Imran Jami, and Shaukat Wasi; *International Journal of Advanced Computer Science and Applications*, volume 10, number 9, pages 546–552, 2019. (WoS ESCI indexed). DOI: 10.14569/IJACSA.2019.0100972.
8. "Developing Disease Classification System based on Keyword Extraction and Supervised Learning". Muhammad Suffian, [Muhammad Yaseen Khan](#), and Shaukat Wasi; *International Journal of Advanced Computer Science and Applications*, volume 9, number 9, pages 599–605, 2018. (WoS ESCI indexed). DOI: 10.14569/IJACSA.2018.090976.

Refereed Book Chapters

9. "Pseudo Transfer Learning by Exploiting Monolingual Corpus: An Experiment on Roman Urdu Transliteration". [Muhammad Yaseen Khan](#) and Tafseer Ahmed; in: Bajwa I., Sibalija T., Jawawi D. (eds) *Intelligent Technologies and Applications. INTAP 2019. Communications in Computer and Information Science*, vol 1198. Springer, Singapore. DOI 10.1007/978-981-15-5232-8_36.
10. "Towards a Generic Approach for PoS Tag-wise Lexical Similarity of Languages". Muhammad Suffian Nizami, [Muhammad Yaseen Khan](#) and Tafseer Ahmed; in: Bajwa I., Sibalija T., Jawawi D. (eds) *Intelligent Technologies and Applications. INTAP 2019. Communications in Computer and Information Science*, vol 1198. Springer, Singapore. DOI 10.1007/978-981-15-5232-8_42.

Refereed Conference Publications

11. "Sign Language Translation in Urdu/Hindi Through Microsoft Kinect". Sohail, [Muhammad Yaseen Khan](#), Nisar Ahmed Siddiqui, and Kiran; in *Proceedings of the 3rd Pak-Turk Conference On Emerging Technologies*, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Pakistan. June. 2020. DOI: 10.1088/1757-899x/899/1/012016.
12. "Urdu Sentiment Corpus (v1.0): Linguistic Exploration and Visualization of Labeled Dataset for Urdu Sentiment Analysis". [Muhammad Yaseen Khan](#) and Suffian Nizami; in *Proceedings of the 2nd IEEE International Conference On Information Science & Communication Technology*, Karachi, Pakistan. Feb. 2020. DOI: 10.1109/ICISCT49550.2020.9080043.
13. "Practices for Achieving Accuracy in Software Costing and Estimation". Mohammad Ayub Latif, [Muhammad Yaseen Khan](#), and Kashif Bashir; in *Proceedings of the 2nd International Conference on Computing & Information Sciences*, Karachi Institute of Economics and Technology, Karachi, Pakistan. Nov. 2018.
14. "Drive pattern analyzer using smartphones". Shumaila Mahboob, Kanwal Iqbal, Hira Farman, [Muhammad Yaseen Khan](#), and Shaukat Wasi; in *Proceedings of the 13th IEEE International Conference on Emerging Technologies*, Capital University of Science and Technology, Islamabad, Pakistan. IEEE, pages 1-5, 2017. DOI: 10.1109/ICET.2017.8281659
15. "Harnessing English Sentiment Lexicons for Polarity Detection in Urdu Tweets: A Baseline Approach". [Muhammad Yaseen Khan](#), Shah Muhammad Emaduddin, and Khurum Nazir Junejo; in *Proceedings of the 11th IEEE International Conference on Semantic Computing*, San Diego, USA. IEEE, pages. 242-24, 2017. DOI: 10.1109/ICSC.2017.68

Unpublished Work

15. "Discovering Lexical Similarity Through Articulatory Feature-based Phonetic Edit Distance". Tafseer Ahmed, Muhammad Suffian Nizami, [Muhammad Yaseen Khan](#); <https://arxiv.org/abs/2008.06865>
16. "Predicting Good Dictionary Examples with Distant Supervision and Machine Learning Techniques". [Muhammad Yaseen Khan](#), Abdul Qayoom Bhatti, and Tafseer Ahmed; *experimentation completed; in draft*.
17. "Developing Sentiment Lexicon and Labelled Corpus for Urdu Sentiment Analysis". [Muhammad Yaseen Khan](#), Rafia Zaheer; *experimentation completed; in draft*.
18. "Developing Emotion Lexicon and Labelled Corpus for Urdu Emotion Detection". [Muhammad Yaseen Khan](#), Thereem Khalid; *experimentation completed; in draft*.
19. "Urdu Named-Entity Recognition System". Tafseer Ahmed, [Muhammad Yaseen Khan](#), and Suffian Nizami; *experimentation completed; in draft*.
20. "LSTM-based Urdu Word Segmentation". [Muhammad Yaseen Khan](#) and Tafseer Ahmed; *experimentation completed; in draft*.

21. “Depressive Sentiment Analysis in Microblogs: A Machine Learning Approach”. Shah Muhammad Emaduddin, Kamran Ahsan, [Muhammad Yaseen Khan](#), and Sohaib Ahmed; *under review*.
22. “Efficient Recipe Recommendation System Using Signature-Trees”. Ali Sohani, [Muhammad Yaseen Khan](#), and Rafiullah Afridi; *accepted with revision*.
23. “Exploiting Patient Descriptions of Top-20 Diseases in Text Classification for Evidence-based Medicine”. Samiullah Shaikh and [Muhammad Yaseen Khan](#); *in draft*.

Conference and Workshop Attendee

1. 7th Intl. Conf. on Language and Technology, CLE-UET, Lahore (2020)
2. 2nd Intl. Conf. on Information Science & Communication Technology, UoK, Karachi (2020)
3. 2nd Intl. Conf. on Intelligent Technologies and Applications, IUB, Bhahawalpur (2019)
4. 1st Intl. Conf. on Data Science, MAJU, Karachi (2019)
5. 2nd Intl. Conf. on Computing and Information Sciences, KIET, Karachi (2018)
6. HEC-NCEAC 3-Day Workshop on Data Science, Islamabad (2017)

Current Students

1. Alvina Naveed (MS Thesis - SE); “Exploiting Bug Narrations for Predictive Analysis on Bug Severity”. Expected Fall 2020. Co-supervised by Dr Syed Ghanzanfer Monir.
2. Abdul Basit (MS Thesis - CS); “Human motion tracking for gait analysis through smart phones and machine learning approaches.”. Expected Fall 2020.
3. Abdul Qayoom Bhatti (MS Thesis - CS); “Supervised Learning-based Good Dictionary Examples Recommendation System”. Expected Fall 2020.
4. Mirza Muhammad Muneeb Baig (MS Thesis - CS); “Sarcasm detection in Urdu Text”. Expected Fall 2020.

Graduated Students

1. Muhammad Suffian Nizami (MS Thesis - CS); “Formulating Patient Descriptions to Support Evidence Based Medicine”. Completed May 2018. Co-supervised by Dr Shaukat Wasi.

Awards and Honours

1. Awardee of Dean’s List 2014 and 2015 during MS.SE, KIET

Research Interests

Natural Language Processing · Information Extraction · Information Retrieval · Affective Computing in Text · Lexical Analysis · Discourse Analysis · Computational Linguistics · Corpus Linguistics · Sociolinguistics · NLP for scared source Indic languages (specifically languages in Perso-Arabic scripts)

Skills

Computing Programming (Python, Java, JavaScript, C#, MySQL) · Typesetting with \LaTeX
Languages English · Urdu · Hindi · Persian/Farsi

References

1. Dr. Tafseer Ahmed Khan, Professor and Head Center for Language Computing, MAJU, Karachi.
2. Dr. Shaukat Wasi, Professor and Dean Faculty of Computing, MAJU, Karachi.
3. Dr. Asim Imdad Wagan, Professor and Dean Faculty of Engineering, MAJU, Karachi.
4. Dr. Khurum Nazir Junejo, Head of R&D at Love for Data.
5. Dr. Muhammad Khalid Khan, Professor and Director CoCIS, PAF-KIET, Karachi.