

# 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. S. 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. S. 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. Technical committee, 1st Mohammad Ali Jinnah University International Conference on Computing (MAJICC '21), MAJU, Karachi (2021)
2. Organising committee, 1st International Conference on Data Sciences, MAJU, Karachi (2019)

3. Trainer, 1-Day Workshop on Data Science with Python, UBIT, Karachi University (2018)
4. Trainer, Hands-on Programming with Python, MAJU, Karachi (2018)
5. Trainer, Programming with Python, Combat:18, KIET, Karachi (2018)
6. Coach, ICPC 2015 @ NUCES-FAST, Lahore (2015)
7. Faculty head, CodeJam, PAF-KIET, Karachi (2015)
8. Faculty head, Combat: The Computing Battle 2017, KIET, Karachi (2014)
9. 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 covered into book chapters or journal articles are counted once.

### Refereed Journal Articles

1. “Automated Prediction of Good Dictionary EXamples (GDEX): A Comprehensive Experiment with Distant Supervision, Machine Learning and Word Embedding-based Deep Learning Techniques”. [M. Y. Khan](#), A. Qayoom, M. S. Nizami, M. S. Siddiqui, S. Wasi, and S. M. K. R. Raazi; *Complexity*, volume 11, number 11, pages 749–755, 2020. (WoS **SCIE** indexed /JCR-IF: **2.83**)
2. “Hindustani or Hindi vs. Urdu: A Computational Approach for the Exploration of Similarities Under Phonetic Aspects”. M. S. Nizami, T. Ahmed, and [M. Y. Khan](#); *International Journal of Advanced Computer Science and Applications*, volume 11, number 11, pages 749–755, 2020. DOI: [10.14569/IJACSA.2020.0111191](#), (WoS **ESCI** indexed).
3. “Edit Distance-based Search Approach for Retrieving Element-wise Prosody in Hindi/Urdu Poetry”. [M. Y. Khan](#), M. A. Rao, S. Wasi, T. A. Minai; *Indian Journal of Science and Technology*, volume 13, number 39, pages 4189–420. , 2020. (MJL indexed). DOI: [10.17485/IJST/v13i39.1489](#)
4. “Exerting 2D-Space of Sentiment Lexicons with Machine Learning Techniques: A Hybrid Approach for Sentiment Analysis”. [M. Y. Khan](#) and K. N. 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](#).
5. “Sentiment Analysis of Product Reviews in the Absence of Labelled Data using Supervised Learning Approaches”. W. Muhammad, M. Mushtaq, K. N. Junejo, and [M. Y. Khan](#); *Malaysian Journal of Computer Science*, volume 32, number 2, pages 118–132, 2020. (WoS **SCI** indexed). DOI: [10.22452/MJCS.vol33no2.3](#). JCR-IF: **0.7**
6. “Depression Detection using Naïve Bayes: Text based approach”. S. M. Emad, K. Ahsan, [M. Y. Khan](#), S. Ahmed, R. Ullah, K. Rizwan, and F. Shafiq; *Interciencia*, volume 44, number 11, pages 174–182, 2020. (WoS JCR-IF: **SCIE** indexed). JCR-IF: **0.26**.

7. "On Extraction of Event Information from Social Text Stream: An Unpretentious NLP Solution". K. Iqbal, [M. Y. Khan](#), S. Wasi, S. Mahboob, and T. Ahmed; *International Journal of Computer Science and Network Security*, volume 19, number 9, pages 121–131, 2019 (WoS ESCI indexed).
8. "A Novel Approach For Ontology-Driven Information Retrieving Chatbot For Fashion Brands". A. Nazir, [M. Y. Khan](#), T. Ahmed, S. I. Jami, and S. 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](#).
9. "Developing Disease Classification System based on Keyword Extraction and Supervised Learning". M. Suffian, [M. Y. Khan](#), and S. 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

10. "Pseudo Transfer Learning by Exploiting Monolingual Corpus: An Experiment on Roman Urdu Transliteration". [M. Y. Khan](#) and T. 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](#).
11. "Towards a Generic Approach for PoS Tag-wise Lexical Similarity of Languages". M. S. Nizami, [M. Y. Khan](#) and T. 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

12. "Classic Fairy Tales / Children Stories of the World: Corpus Organization and Text Analytics". [M. Y. Khan](#) and M. U. T. Butt; in *Proceedings of the 7th International Conference of Linguistic Association of Pakistan (ICLAP '21)*, Sargodha, Pakistan. Nov. 2021.
13. "Using Patient Descriptions of Top-20 Diseases in Text Classification for Evidence-based Medicine". Samiullah Shaikh, [M. Y. Khan](#), and M. S. Nizami; in *Proceedings of the 1st Mohammad Ali Jinnah university International Conference on Computing (MAJICC '21)*, Karachi, Pakistan. July. 2021. DOI: [10.1109/MAJICC53071.2021.9526252](#).
14. "Processing Fire Detection, Face Recognition, and Head Counting for an IoT-based Application in Emergency Exit". K. Badr, K. Iqbal, S. Mahboob, and [M. Y. Khan](#); in *Proceedings of the 1st Mohammad Ali Jinnah university International Conference on Computing (MAJICC '21)*, Karachi, Pakistan. July. 2021. [10.1109/MAJICC53071.2021.9526261](#).
15. "Sign Language Translation in Urdu/Hindi Through Microsoft Kinect". Sohail, [M. Y. Khan](#), N. A. 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](#).
16. "Urdu Sentiment Corpus (v1.0): Linguistic Exploration and Visualization of Labeled Dataset for Urdu Sentiment Analysis". [M. Y. Khan](#) and M. S. Nizami; in *Proceedings of the 2nd IEEE International Conference On Information Science & Communication Technology*, Karachi, Pakistan. Feb. 2020. DOI: [10.1109/ICISCT49550.2020.9080043](#).
17. "Practices for Achieving Accuracy in Software Costing and Estimation". M. A. Latif, [M. Y. Khan](#), and K. Bashir; in *Proceedings of the 2nd International Conference on Computing & Information Sciences*, Karachi Institute of Economics and Technology, Karachi, Pakistan. Nov. 2018.
18. "Drive pattern analyzer using smartphones". S. Mahboob, K. Iqbal, H. Farman, [M. Y. Khan](#), and S. 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](#).
19. "Harnessing English Sentiment Lexicons for Polarity Detection in Urdu Tweets: A Baseline Approach". [M. Y. Khan](#), S. M. Emaduddin, and K. N. 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

20. "Discovering Lexical Similarity Through Articulatory Feature-based Phonetic Edit Distance". T. Ahmed, M. S. Nizami, [M. Y. Khan](#); <https://arxiv.org/abs/2008.06865>.

21. “Developing Sentiment Lexicon and Labelled Corpus for Urdu Sentiment Analysis”. [Muhammad Yaseen Khan](#), R. Zaheer; *experimentation completed; in draft*.
22. “Developing Emotion Lexicon and Labelled Corpus for Urdu Emotion Detection”. [Muhammad Yaseen Khan](#), T. Khalid; *experimentation completed; in draft*.
23. “Urdu Named-Entity Recognition System”. T. Ahmed, [M. Y. Khan](#), and S. Nizami; *experimentation completed; in draft*.
24. “LSTM-based Urdu Word Segmentation”. [M. Y. Khan](#) and T. Ahmed; *experimentation completed; in draft*.
25. “Depressive Sentiment Analysis in Microblogs: A Machine Learning Approach”. S. M. Emaduddin, K. Ahsan, [M. Y. Khan](#), and S. Ahmed; *under review*.
26. “Efficient Recipe Recommendation System Using Signature-Trees”. A. Sohani, [M. Y. Khan](#), and R. Ullah; *accepted with revision*.

## Conference and Workshop Attendee

1. 7th Intl. Conf. of Linguistics of Pakistan (ICLAP), University of Sargodha, Pakistan (2021)
2. 1st Mohammad Ali Jinnah University Conf. on Computing (MAJICC), Karachi, Pakistan (2021)
3. 7th Intl. Conf. on Language and Technology (CLT), CLE-UET, Lahore (2020)
4. 2nd Intl. Conf. on Information Science & Communication Technology (ICISCT), UoK, Karachi, Pakistan (2020)
5. 2nd Intl. Conf. on Intelligent Technologies and Applications (INTAP), IUB, Bhawalpur (2019)
6. 1st Intl. Conf. on Data Science, MAJU, Karachi, Pakistan (2019)
7. 2nd Intl. Conf. on Computing and Information Sciences, KIET, Karachi, Pakistan (2018)
8. HEC-NCEAC 3-Day Workshop on Data Science, Islamabad (2017)

## Graduated Students

1. M. S. Nizami (MS Thesis - CS); “Formulating Patient Descriptions to Support Evidence Based Medicine”. Spring 2018. Co-supervised by Dr S. Wasi.
2. Alvina Naveed (MS Thesis - SE); “Exploiting Bug Narrations for Predictive Analysis on Bug Severity”. Fall 2020.
3. Abdul Basit (MS Thesis - CS); “Human motion tracking for gait analysis through smart phones and machine learning approaches”. Fall 2020.
4. Abdul Qayoom Bhatti (MS Thesis - CS); “Supervised Learning-based Good Dictionary Examples Recommendation System”. Fall 2020.
5. Mirza Muhammad Muneeb Baig (MS Thesis - CS); “Sarcasm detection in Urdu Text”. Fall 2020.

## 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 scarce source Indic languages (specifically languages in Perso-Arabic scripts)

## Skills

Computing . . . . . Programming (Python, Java, JavaScript, C#, MySQL) · Typesetting with  $\text{\LaTeX}$   
 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.