Dr Muhammad Asif Naeem

Curriculum Vitae

Address: Department of Data Science & Artificial Intelligence,
National University of Computer and Emerging Sciences
3 A.K. Brohi Road, H-11/4, Islamabad, Pakistan.

Cell: +92 335 9445737, **E-mail:** asif.naeem@nu.edu.pk
Homepage: `http://isb.nu.edu.pk/asif/`
Research Lab: `http://isb.nu.edu.pk/datainsight/`
Google Scholar: `https://shorturl.at/vUQ6A`
LinkedIn: `https://shorturl.at/yiibu`
Nationality: New Zealand

Citations: 2651 | h-index: 23 | i10-index: 53 | Impact Factor: 217

Academic Qualifications

2007 – 2011	**PhD with best thesis award** in Real-time Data Management, Knowledge Management Group, Department of Computer Science, The University of Auckland, New Zealand.
2003 – 2005	**MS with Distinction** in Computer Science, BUITEMS, Pakistan.
1998 – 2000	**M.Sc.** in Computer Science, BZU, Pakistan.
1995 – 1998	**B.Sc.**, Government College Bosan Road Multan, Pakistan.

Fellowships

2017 – 2017	**Fellow of the Higher Education Academy (FHEA)**, UK
2011 – 2012	**Post-doctorate** in the area of Data Stream Management, The University of Auckland, New Zealand.

Research Interests

Generative AI, Large Language Models, Machine Learning, Deep Learning, Data Mining, Data Science, Natural Language Processing, Social Data Analytics, Databases & Data Warehousing.

Academic Experience

2023 – Present	**Director ORIC (Office of Research Innovation & Commercialisation**, National University of Computer & Emerging Science (NUCES), (among top 3 Universities in CS in Pakistan)
2021 – Present	**Director**, Data Insight Research Lab at NUCES
2019 – Present	**Professor**, Department of AI&DS, NUCES
2016 – 2019	**Founding Member**, INTERACT Centre of Technology Excellence at Auckland University of Technology, New Zealand.
2015 – Present	**Founder**, Data Science Research Centre at Auckland University of Technology
2017 – 2019	**Academic Advisor**, School of Engineering, Computer and Mathematical Sciences
2014 – 2019	**Senior Lecturer**, School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology.
2014 – 2019	**Research Member**, Science, Technology, Engineering, and Mathematics Tertiary Education Centre at Auckland University of Technology.

2013 – 2019	**Research Member**, Centre for Artificial Intelligence Research (CAIR) at Auckland University of Technology.
2012 – Present	**Honorary Researcher**, The University of Auckland.
2012 – 2014	**Lecturer**, School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology.
2011 – 2012	**Postdoctoral Fellow**, The University of Auckland.
2010 – 2012	**Research Member**, Knowledge Management Group (KMG) at The University of Auckland.
2012 – 2012	**Visiting Fellow**, Database Systems and Information Management Group at TU Berlin, Germany.
2009 – 2011	**Research Assistant**, Department of Computer Science, The University of Auckland.
2007 – 2011	**PhD Student**, Department of Computer Science, The University of Auckland.
2002 – 2007	**Assistant Professor & Lecture**, Department of Computer Science, Balochistan University of Information Technology, Engineering & Mathematical Sciences (BUITEMS), Pakistan.
2001 – 2002	**IT Incharge**, Government Science College Quetta, Pakistan.

Industry Experience

2024 – present	**Consultant** with APIMatic for project *API Copilot: VS Code Extension*
2024 – present	**Consultant** with New Zealand College of Chiropractic for project *Assistance with final-stage refinements of the BodySync*
2024 – present	**Consultant** with Motive for project *AutoMorph*
2024 – present	**Consultant** with Elixir for project *IntelliDatum*
2023 – 2024	**Consultant** with Upstart Commerce for projects *ImageMart* and *CaptionCraft*

Research Funding

2022	**National Research Program for Universities (NRPU 2022) from Higher Education Commission of Pakistan (PKR.7,108,500)** for project *Make Data Useful from Rubble: An AI-based Approach to Digitize Handwritten Medical Prescriptions and to Build a Recommender System* **(as PI)**.
2020	**Capacity Building Grant for IKMA Lab (PKR.1,375,000) (as Co-PI)**.
2020	**Faculty Research Support Grant (PKR.590,000)** from NUCES for project *A Deep Level Shopping Cart Analysis* **(as PI)**.
2019	**Callaghan Innovation New Zealand (NZ$100,000)** for project *Tolerant Machine Learning for Deficient Training Data* **(as Co-PI)**.
2018	**Callaghan Innovation New Zealand (NZ$100,000)** for project *Narrative extraction and synthesis to make sense of visual media coming from a camera* **(as Co-PI)**.
2017	**Strategic Research Innovation Fund (SRIF) AUT (≈NZ$900,000)** involve in *INTERACT Centre of Technology Excellence Establishment project* **(as PI)**
2017	**Postdoctoral Funding (NZ$160,000)** from AUT for project *Advanced Data Analytics with Applications to Financial Economics* **(as Member)**
2017	**SRIF AUT (NZ$50,000)** for project *Real-time Streaming Analytics*
2016	**Mid-career Research Grant (NZ$25,000)** for project *Bridging the Gap between Real-time Big Data and Persistent Big Data* **(as PI)**
2016	**Faculty Contestable Research Grant (NZ$20,000)** for project *Smart Shopping with Real-time Shopping Cart Analysis* **(as PI)**

2016 **SRIF (NZ$20,000)** for project *Forecasting Frequency and Time Synchronization Fluctuations in Electrical Signals* **(as PI)**

2016 **Orion Precision Driven Health Grant (NZ$5,500)** for project *Post Discharge Application* **(as PI)**

2015 **SRIF (NZ$50,000)** for project *Supporting Real-time Processing of Gaze Data in a Brick-and-Mortar BaM 2.0* **(as PI)**

Distinctions and Awards

2018 **Finalist for VC Research Excellence Award, AUT**

2016 **Dean's Research Excellence Award, AUT**

2016 **Finalist for VC Research Excellence Award, AUT**

2015 **Best Faculty Teaching Award, AUT**

2012 **Best PhD Thesis Award**, Department of Computer Science, The University of Auckland

2012 **In-time PhD Completion Award ($6,000)**, The University of Auckland

2011 **Best Poster Award**, Department of Computer Science, The University of Auckland

2009 **Best Poster Award**, Faculty of Science, The University of Auckland

2006 **Distinction Award** in MS Computer Science

Research Publications

Journal Articles

1. **M. Asif Naeem**, Saad Munir, Warsha Khan, Hamza Khan, and Muhammad Ali Tahir. Understanding handwritten medical documents using transformer based learning. *Soft Computing*, 2025, Accepted.

2. UHWA Hewage, Roopak Sinha, and **M. Asif Naeem**. An accuracy-privacy optimization frameworkconsidering users privacy requirements for datastream mining. *Journal of Big Data*, 2025, Accepted.

3. Adeel Cheema and **M. Asif Naeem**. Covergan: cover photo generation from text story using layout guided gan. *Soft Computing*, pages 1–19, 2025.

4. Widana Kankanamge Darsha Jayamini, Farhaan Mirza, **M. Asif Naeem**, Amy Hai Yan Chan, Andrew Tomlin, Holly Tibble, and Kebede Beyene. Predicting asthma attacks in new zealand using machine learning. *European Respiratory Journal*, 64(suppl 68), 2024.

5. Widana Kankanamge Darsha Jayamini, Farhaan Mirza, Marie-Claire Bidois-Putt, **M. Asif Naeem**, and Amy Hai Yan Chan. Perceptions towards using artificial intelligence and technology for asthma a qualitative exploration of mori views. *European Respiratory Journal*, 64(suppl 68), 2024.

6. Farhana Sharief, Humaira Ijaz, Mohammad Shojafar, and **M. Asif Naeem**. Multi-class imbalanced data handling with concept drift in fog computing: A taxonomy, review, and future directions. **ACM Computing Surveys**, 57(1):1–48, 2024.

7. Adil Majeed, Usama Imtiaz, **M. Asif Naeem**, Muhammad Aleem, Waseem Shahzad, Mirza Omer Beg, and Hasan Mujtaba. Extracting emotion from resource poor language through transfer learning. *Multimedia Tools and Applications*, pages 1–18, 2024.

8. Amy Hai Yan Chan, Braden Te Ao, Christina Baggott, Alana Cavadino, Amber A Eikholt, Matire Harwood, Joanna Hikaka, Dianna Gibbs, Mariana Hudson, Farhaan Mirza, **M. Asif Naeem**, et al. Digipredict: physiological, behavioural and environmental predictors of asthma attacksa prospective observational study using digital markers and artificial intelligencestudy protocol. *BMJ open respiratory research*, 11(1):e002275, 2024.

9. Widana Kankanamge Darsha Jayamini, Farhaan Mirza, **M. Asif Naeem**, and Amy Hai Yan Chan. Investigating machine learning techniques for predicting risk of asthma exacerbations: A systematic review. *Journal of Medical Systems*, 48(1):49, 2024.

10. Widana Kankanamge Darsha Jayamini, Farhaan Mirza, Marie-Claire Bidois-Putt, **M. Asif Naeem**, and Amy Hai Yan Chan. Perceptions toward using artificial intelligence and technology for asthma attack risk prediction: Qualitative exploration of māori views. *JMIR Formative Research*, 8(1):e59811, 2024.

11. Saad Munir and **M. Asif Naeem**. Bil-fand: leveraging ensemble technique for efficient bilingual fake news detection. *International Journal of Machine Learning and Cybernetics*, pages 1–23, 2024.

12. Musa Dildar Ahmed Cheema, Mohammad Daniyal Shaiq, Farhaan Mirza, Ali Kamal, and **M. Asif Naeem**. Adapting multilingual vision language transformers for low-resource urdu optical character recognition (ocr). *PeerJ Computer Science*, 10:e1964, 2024.

13. Benjamin Denham, Edmund MK Lai, Roopak Sinha, and **M. Asif Naeem**. Dynamic quantification with constrained error under unknown general dataset shift. **IEEE Transactions on Knowledge and Data Engineering (TKDE)**, 2024.

14. UHWA Hewage, R Sinha, and **M. Asif Naeem**. Privacy-preserving data (stream) mining techniques and their impact on data mining accuracy: a systematic literature review. *Artificial Intelligence Review*, pages 1–38, 2023.

15. Benjamin Denham, Edmund MK Lai, Roopak Sinha, and **M. Asif Naeem**. Witan: Unsupervised labelling function generation for assisted data programming. **Proceedings of the VLDB Endowment**, 15(11):2334 – 2347, 2022.

16. Muhammad Usama, **M. Asif Naeem**, and Farhaan Mirza. Multi-class skin lesions classification using deep features. *Sensors*, 22(21):8311, 2022.

17. Widana Kankanamge Darsha Jayamini, Farhaan Mirza, **M. Asif Naeem**, and Amy Hai Yan Chan. State of asthma-related hospital admissions in new zealand and predicting length of stay using machine learning. *Applied Sciences*, 12(19), 2022.

18. **M. Asif Naeem**, Wasiullah Waqar, Farhaan Mirza, and Ali Tahir. Tinylfu-based semi-stream cache join for near-real-time data warehousing. *Soft Computing*, 2022.

19. Rashi Bhalla, Russel Pears, **M. Asif Naeem**, and Farhaan Mirza. Novel method for optimizing performance in resource constrained distributed data streams. *Applied Intelligence*, pages 1–19, 2022.

20. U.H.W.A. Hewage, Russel Pears, and **M. Asif Naeem**. Optimizing the Trade-off Between Classification Accuracy and Data Privacy in the Area of Data Stream Mining. *International Journal of Artificial Intelligence*, 1(1):147–167, 2022.

21. Sana Aurangzeb, Haris Anwar, **M. Asif Naeem**, and Muhammad Aleem. Bigrc-eml: big-data based ransomware classification using ensemble machine learning. *Cluster Computing*, pages 1–18, 2022.

22. Herman Masindano Wandabwa, **M. Asif Naeem**, Farhaan Mirza, and Russel Pears. Multi-interest semantic changes over time in short-text microblogs. **Knowledge-Based Systems**, 228:107249, 2021.

23. Farrukh Zahid, Ali Tahir, Habib Ullah Khan, and **M. Asif Naeem**. Wind farms selection using geospatial technologies and energy generation capacity in gwadar. *Energy Reports*, 7:5857–5870, 2021.

24. Muhammad Daud Kamal, Ali Tahir, Muhammad Babar Kamal, and **M. Asif Naeem** Naeem. Future location prediction for emergency vehicles using big data: A case study of healthcare engineering. *Journal of Healthcare Engineering*, 2020, 2020.

25. Muhammad Daud Kamal, Ali Tahir, Muhammad Babar Kamal, Faisal Moeen, and **M. Asif Naeem**. A survey for the ranking of trajectory prediction algorithms on ubiquitous wireless sensors. *Sensors*, 20(22):6495, 2020.

26. Herman Masindano Wandabwa, **M. Asif Naeem**, Farhaan Mirza, and Russel Pears. Topical affinity in short text microblogs. **Information Systems**, page 101662, 2020.

27. **M. Asif Naeem**, Farhaan Mirza, Habib Ullah Khan, David Sundaram, Noreen Jamil, and Gerald Weber. Big data velocity management-from stream to warehouse via high performance memory optimised index join. *IEEE Access*, 8:195370–195384, 2020.

28. **M. Asif Naeem**, Habib Ullah Khan, Saad Aslam, and Noreen Jamil. Parallelisation of a cache-based stream-relation join for a near-real-time data warehouse. *Electronics*, 9(8):1299, 2020.

29. Yiwei Feng, **M. Asif Naeem**, Farhaan Mirza, and Ali Tahir. Reply using past replies - a deep learning-based e-mail client. *Electronics*, 9(9):1353, 2020.

30. Asif Mansoor, Muhammad Waleed Usman, Noreen Jamil, and **M. Asif Naeem**. Deep learning algorithm for brain-computer interface. *Scientific Programming*, 2020, 2020.

31. Benjamin Denham, Russel Pears, and **M. Asif Naeem**. Enhancing random projection with independent and cumulative additive noise for privacy-preserving data stream mining. **Expert Systems with Applications**, page 113380, 2020, (*Q1 in SCImago,IF=4.577*).

32. Amira Khattak, Noreen Jamil, **M. Asif Naeem**, Farhaan Mirza, et al. Data analytics in mental healthcare. *Scientific Programming*, 2020, 2020.

33. Mohammed Alkorbi, Noreen Jamil, **M. Asif Naeem**, Farhaan Mirza, et al. Evaluating encryption algorithms for sensitive data using different storage devices. *Scientific Programming*, 2020, IF=1.289.

34. Benjamin James Denham, Russel Pears, and **M. Asif Naeem**. HDSM: A distributed data mining approach to classifying vertically distributed data streams. **Knowledge-Based Systems**, pages 105–114, 2020, (*Q1 in SCImago,IF=5.358*).

35. **M. Asif Naeem**, Erum Mehmood, and M. G. Abbas Malik. Optimizing semi-stream cachejoin for near-real-time data warehousing. *Journal of Database Management*, 2020, (*Q2 in SCImago, IF=0.12*).

36. Adi Darliansyah, Herman Masindano Wandabwa, **M. Asif Naeem**, Farhaan Mirza, and Russel Pears. Sentipede: A smart system for sentiment-based personality detection from short texts. *Journal of Universal Computer Science*, 2019, (*Q2 in SCImago,IF=0.77*).

37. Huy Vuong Nguyen, **M. Asif Naeem**, Nuttanan Wichitaksorn, and Russel Pears. A smart system for short-term price prediction using time series models. *Journal of Computer and Electrical Engineering*, 2019, (*IF=2.33*).

38. **M. Asif Naeem**. Optimisation and extension of stream-relation joins. *The International Journal of Information Technology & Decision Making*, 2019, (*ranked Q1 in SCImago, IF=2.86*).

39. M. Safdar Munir, Imran Sarwar Bajwa, **M. Asif Naeem**, and Bushra Ramzan. Design and implementation of an iot system for smart energy consumption and smart irrigation in tunnel farming. *Energies*, 11(12), 2018, (*ranked Q1 in SCImago, IF=2.70*).

40. **M. Asif Naeem**, Aftab A Mughal, Christof Lutteroth, and Gerald Weber. A smart email client prototype for effective reuse of past replies. *IEEE Access*, 6(1):69453–69471, 2018, (*ranked Q1 in SCImago, IF=4.08*).

41. Herman Wandabwa, **M. Asif Naeem**, Farhaan Mirza, and Russel Pears. A metamodel enabled approach for discovery of coherent topics in short text microblogs. *IEEE Access*, 6(1):65582–65593, 2018, (*ranked Q1 in SCImago, IF=4.08*).

42. **M. Asif Naeem**, Christof Lutteroth, and Gerald Weber. A memory-optimal many-to-many semi-stream join. **Distributed and Parallel Databases**, pages 1–27, 2018, (*ranked A in CORE, IF=1.27, Q2 in SCImago*).

43. Hoa Nguyen, Farhaan Mirza, **M. Asif Naeem**, and Mansoor Baig. Falls management framework for supporting an independent lifestyle for older adults a systematic review. *Aging Clinical and Experimental Research*, pages 1–12, 2018, (*IF=2.33, Q2 in SCImago*).

44. Noreen Jamil, Farhaan Mirza, **M. Asif Naeem**, and Nilufar Baghaei. A refinement of an iterative orthogonal projection method. **Journal of Computational and Applied Mathematics**, 341:31–41, 2018, (*ranked A in ARC, IF=2.014, Q1 in SCImago*).

45. Chamari I. Kithulgoda, Russel Pears, and **M. Asif Naeem**. The incremental fourier classifier: Leveraging the discrete fourier transform for classifying high speed data streams. **Expert Systems with Applications**, 97:1–17, 2018, (*IF=4.57, Q1 in SCImago*).

46. **M. Asif Naeem**, Gillian Dobbie, Christof Lutteroth, and Gerald Weber. Skewed distributions in semi-stream joins: How much can caching help? **Information Systems**, 64:63–74, 2017, (*ranked A* in CORE, IF=2.67, Q1 in SCImago*).

47. Imran Sarwar Bajwa, Fatima Karim, **M. Asif Naeem**, and Riaz Ul Amin. A semi supervised approach for catchphrase classification in legal text documents. *Journal of Computers*, 12(5):451–461, 2017, (*IF=0.78, Q3 in SCImago*).

48. Rivindu Perera, Parma Nand, and **M. Asif Naeem**. Utilizing typed dependency subtree patterns for answer sentence generation in question answering systems. *Progress in Artificial Intelligence*, 6(2):105–119, 2017, (*Q2 in SCImago, IF=0.5*).

49. Imran Sarwar Bajwa, N. Mamoona Asghar, and **M. Asif Naeem**. Automated detection of early tropical cyclones formation in satellite images. *Pakistan Journal of Science*, 69(4), 2017, (*IF=0.27, Q3 in SCImago*).

50. Imran Sarwar Bajwa, N. Mamoona Asghar, and **M. Asif Naeem**. Learning based improved seeded region growing algorithm for brain tumor identification. *Proceeding of Pakistan Academy of Science*, 54(2), 2017, (*IF=0.13, Q4 in SCImago*).

51. Imran Sarwar Bajwa, Nadeem Sarwar, and **M. Asif Naeem**. Generating EXPRESS data models from sbvr. *Proceeding of Pakistan Academy of Sciences*, 53(4):381–389, 2016, (*IF=0.13, Q4 in SCImago*).

52. M. Shahzad Akram, Imran Sarwar Bajwa, and **M. Asif Naeem**. A supervised approach for semantic annotation of entities in text. *Science International (Lahore)*, 28(1):237–241, 2016, (*IF=1.36*).

53. Imran Sarwar Bajwa and **M. Asif Naeem**. A CBR based automated approach for reusing business process models. *Science International (Lahore)*, 28(1):227–232, 2016, (*IF=1.36*).

54. **M. Asif Naeem**, Imran Sarwar Bajwa, and Noreen Jamil. A cached-based stream-relation join operator for semi-stream data processing. *International Journal of Data Warehousing and Mining (IJDWM)*, 12(3):14–31, 2016, (*IF=0.20, Q2 in SCImago*).

55. Noreen Jamil, Johannes Müller, **M. Asif Naeem**, Christof Lutteroth, and Gerald Weber. Extending linear relaxation for non-square matrices and soft constraints. **Journal of Computational and Applied Mathematics**, 308:346–360, 2016, (*ranked A in ARC, IF=2.01, Q1 in SCImago*).

56. **M. Asif Naeem**, Gillian Dobbie, and Gerald Weber. Efficient processing of streaming updates with archived master data in near-real-time data warehousing. **Knowledge and information systems (KAIS)**, 40(3):615–637, 2014, (*IF=2.40, Q1 in SCImago*).

57. **M. Asif Naeem** and Noreen Jamil. An efficient stream-based join to process end user transactions in real-time data warehousing. *Journal of Digital Information Management*, 12(3):201–215, 2014, (*IF=0.13, Q4 in SCImago*).

58. Noreen Jamil and **M. Asif Naeem**. Speeding up sor and kaczmarz for constraint-based guis with a warm-start strategy. *Journal of Multimedia Processing and Technologies (JMPT)*, 4(3):179–188, 2013.

59. **M. Asif Naeem**, Gillian Dobbie, and Gerald Weber. Hybridjoin for near-real-time data warehousing. *International Journal of Data Warehousing and Mining (IJDWM)*, 2011, (*IF=0.20, Q2 in SCImago*).

60. Imran Sarwar Bajwa, Ahsan Ali Chaudhri, and **M. Asif Naeem**. Processing large data sets using a cluster computing framework. *Australian Journal of Basic and Applied Sciences (AJBAS)*, 5(6):1614–1618, 2011, (*IF=0.126, Q4 in SCImago*).

61. **M. Asif Naeem** and Noreen Jamil. A web smart space framework for intelligent search engines. *International Journal of Emerging Sciences*, 1(1):1, 2011.

Conference Papers

1. Eisha Rehan, Shahwaiz Memon, Muhsin Raza, M Asif Naeem, and Fazeel Nadeem Bhatti. A light-weight image to description model using knowledge distillation. In *2024 2nd International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings)*, pages 1–7. IEEE, 2024.

2. Javeria Zia, Usman Habib, and **M. Asif Naeem**. Extraction and classification of medicines from handwritten medical prescriptions. In *2023 18th International Conference on Emerging Technologies (ICET)*, pages 104–109. IEEE, 2023.

3. Iqra Ali and **M. Asif Naeem**. Identifying and profiling user interest over time using social data. In *2022 24th International Multitopic Conference (INMIC)*, pages 1–6. IEEE, 2022.

4. Rida Tahir and **M. Asif Naeem**. A machine learning based approach to identify user interests from social data. In *2022 24th International Multitopic Conference (INMIC)*, pages 1–6. IEEE, 2022.

5. Muhammad Farjad Ali Raza and **M. Asif Naeem**. Saraiki language word prediction and spell correction framework. In *2022 24th International Multitopic Conference (INMIC)*, pages 1–6. IEEE, 2022.

6. Benjamin Denham, Edmund MK Lai, Roopak Sinha, and **M. Asif Naeem**. Witan: Unsupervised labelling function generation for assisted data programming. In **Proceedings of 48th International Conference on Very Large Databases (VLDB 2022)**. VLDB Endowment Inc, 2022, (*ranked A* in CORE*).

7. Benjamin Denham, Edmund Lai, Roopak Sinha, and **M. Asif Naeem**. Gain-some-lose-some: Reliable quantification under general dataset shift. In **21th IEEE International Conference on Data Mining (ICDM 2021)**. IEEE, 2021, (*ranked A* in CORE*).

8. Benjamin Denham, Russel Pears, and **M. Asif Naeem**. Null-labelling: A generic approach for learning in the presence of class noise. In **20th IEEE International Conference on Data Mining (ICDM 2020)**. IEEE, 2020, (*ranked A* in CORE*).

9. Herman Wandabwa, **M. Asif Naeem**, Farhaan Mirza, Russel Pears, and Andy Nguyen. Multi-interest user profiling in short text microblogs. In **15th International Conference on Design Science Research in Information Systems and Technology (DESRIST'20)**. Springer LNCS, 2020, (*ranked A in CORE*).

10. Herman Wandabwa, **M. Asif Naeem**, Farhaan Mirza, and Russel Pears. Follow-back recommendations for sports bettors: A twitter-based approach. In **53rd Hawaii International Conference on System Sciences (HICCS)**. Association for Information Systems IEEE Computer Society Press, 2020, (*ranked A in CORE*).

11. Nawal Chanane, Farhaan Mirza, and **M. Asif Naeem**. Co-designing a medication notification application with multi-channel reminders. In **31st Australasian Conference on Information Systems (ACIS'20)**. Association for Information Systems, 2020, (*ranked A in CORE*).

12. Nawal Chanane, Farhaan Mirza, and **M. Asif Naeem**. Insights of medication adherence management: A qualitative study with healthcare professionals and technology designers. In **30th Australasian Conference on Information Systems (ACIS'19)**. Association for Information Systems, 2019, (*ranked A in CORE*).

13. Christopher J. Rapson, Boon-Chong Seet, **M. Asif Naeem**, Jeong Eun Lee, and Reinhard Klette. A performance comparison of deep learning methods for real-time localisation of vehicle lights in video frames. In *IEEE Intelligent Transportation Systems Conference (ITSC19)*. IEEE, 2019.

14. Adi Darliansyah, Herman Masindano Wandabwa, **M. Asif Naeem**, Farhaan Mirza, and Russel Pears. Long-term trends in public sentiment in indian demonetisation policy. In *Intelligent Technologies and Applications*, pages 65–75, Singapore, 2019. Springer.

15. Shraddha Nayak, Md. Akbar Hossain, Farhaan Mirza, **M. Asif Naeem**, and Noreen Jamil. E-brace: A secure electronic health record access method in medical emergency. In *Intelligent Technologies and Applications*, pages 16–27, Singapore, 2019. Springer.

16. **M. Asif Naeem**, Omer Aziz, and Noreen Jamil. Optimising HYBRIDJOIN to process semi-stream data in near-real-time data warehousing. In *CONF-IRM 2019 Proceedings*. Association for Information Systems, 2019.

17. Herman Wandabwa, **M. Asif Naeem**, Farhaan Mirza, and Russel Pears. Topical expressivity in short texts. In *CONF-IRM 2019 Proceedings*. Association for Information Systems, 2019.

18. Christopher J Rapson, Boon-Chong Seet, **M. Asif Naeem**, Jeong Eun Lee, Mahmoud Al-Sarayreh, and Reinhard Klette. Reducing the pain: A novel tool for efficient ground-truth labelling in images. In **Proceedings of 33rd IEEE Conference on Image and Vision Computing New Zealand (IVCNZ)**. IEEE, 2018, (*ranked B in CORE*).

19. Akbar Hossain, Farhaan Mirza, **M. Asif Naeem**, and Jairo Gutierrez. A crowd sourced framework for neighbour assisted medical emergency system. In **2017 27th International Conference on Telecommunication Networks and Applications (ITNAC)**, pages 1–6. IEEE, 2017, (*ranked B in CORE*).

20. Hoa Nguyen, Farhaan Mirza, **M. Asif Naeem**, and Mirza Mansoor Baig. Detecting falls using a wearable accelerometer motion sensor. In *Proceedings of 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MOBIQUITOUS 2017)*, 2017, (*ranked A in CORE*).

21. Herman Wandabwa, **M. Asif Naeem**, and Farhaan Mirza. Aspect of blame in tweets: A deep recurrent neural network approach. In **26th International Conference on World Wide Web Companion**, pages 1423–1424. International World Wide Web Conferences Steering Committee, 2017, (*ranked A* in CORE*).

22. **M. Asif Naeem**, Kim Tung Nguyen, and Gerald Weber. A multi-way semi-stream join for a near-real-time data warehouse. In **28th Australasian Database Conference**. Springer, 2017, (*ranked B in CORE*).

23. Ali Haider Hussein Ghazala, **M. Asif Naeem**, Farhaan Mirza, and Noreen Jamil. Uncovering useful patterns in shopping cart data. In **21st International Conference on Computer Supported Cooperative Work in Design (CSCWD 2017)**. IEEE, March 2017, (*ranked B in CORE*).

24. Herman Wandabwa, **M. Asif Naeem**, and Farhaan Mirza. Document level semantic comprehension of noisy text streams via convolutional neural networks. In **21st International Conference on Computer Supported Cooperative Work in Design (CSCWD 2017)**. IEEE, March 2017, (*ranked B in CORE*).

25. Hoa Hong Nguyen, Farhaan Mirza, **M. Asif Naeem**, and Minh Nguyen. A review on iot healthcare monitoring applications and a vision for transforming sensor data into real-time clinical feedback. In *21st International Conference on Computer Supported Cooperative Work in Design (CSCWD 2017)*. IEEE, March 2017, (*ranked B in CORE*).

26. Nawal Chanane, Farhaan Mirza, **M. Asif Naeem**, and Asfhaan Mirza. Acceptance of technology-driven interventions for improving medication adherence. In *Future Network Systems and Security (FNSS 2017)*. Springer, 2017.

27. **M. Asif Naeem**, Gerald Weber, and Christof Lutteroth. Optimising queue-based semi-stream joins by introducing a queue of frequent pages. In **27th Australasian Database Conference**, pages 407–418. Springer, 2016, (*ranked B in CORE*).

28. Hao Gao, **M. Asif Naeem**, Christof Lutteroth, and Gerald Weber. S3J: A parallel semi-stream similarity join. In *18th International Workshop on Data Warehousing and OLAP*, pages 49–57. ACM, 2015, (*ranked B in CORE*).

29. **M. Asif Naeem**, Imran Sarwar Bajwa, and Noreen Jamil. A cache-based semi-stream join to deal with unmatched stream data. In *26th Australasian Database Conference*, pages 54–65. Springer, 2015, (*ranked B in CORE*).

30. **M. Asif Naeem**, Imran Sarwar Bajwa, and Noreen Jamil. A cached-based approach to enrich stream data with master data. In *10th International Conference on Digital Information Management (ICDIM), 2015*, pages 57–62. IEEE, 2015.

31. **M. Asif Naeem**, Gerald Weber, Christof Lutteroth, and Gillian Dobbie. Optimizing queue-based semi-stream joins with indexed master data. In *16th International Conference on Data Warehousing and Knowledge Discovery*, pages 171–182. Springer, 2014, (*ranked B in CORE*).

32. **M. Asif Naeem**. A caching approach to process stream data in data warehouse. In *9th International Conference on Digital Information Management (ICDIM), 2014*, pages 162–167. IEEE, 2014.

33. Aqsa Mahmood, Kiran Qazi, Imran Sarwar Bajwa, and **M. Asif Naeem**. Natural language processing based interpretation of skewed graphs. In *International Conference on advances in Computing, Communications and Informatics (ICACCI)*, pages 2700–2704. IEEE, 2014.

34. Shabana Ramzan, Imran Sarwar Bajwa, Ikram Ul Haq, and **M. Asif Naeem**. A model transformation from NL to SBVR. In *9th International Conference on Digital Information Management (ICDIM)*, pages 220–225. IEEE, 2014.

35. **M. Asif Naeem**, Gerald Weber, Gillian Dobbie, and Christof Lutteroth. A generic front-stage for semi-stream processing. In **22nd International Conference on Information & Knowledge Management**, pages 769–774. ACM, 2013, (*ranked A in CORE*).

36. **M. Asif Naeem**, Gerald Weber, Gillian Dobbie, and Christof Lutteroth. SSCJ: A semi-stream cache join using a front-stage cache module. In *15th International Conference on Data Warehousing and Knowledge Discovery*, pages 236–247. Springer, 2013, (*ranked B in CORE*).

37. **M. Asif Naeem**. Efficient processing of semi-stream data. In *8th International Conference on Digital Information Management (ICDIM)*, pages 7–10. IEEE, 2013.

38. **M. Asif Naeem**. Tuned X-HYBRIDJOIN for near-real-time data warehousing. In *15th Asia-Pacific Web Conference (APWeb)*, pages 494–505. Springer, 2013.

39. **M. Asif Naeem**. A robust join operator to process streaming data in real-time data warehousing. In *8th International Conference on Digital Information Management (ICDIM)*, pages 119–124. IEEE, 2013.

40. **M. Asif Naeem**, Gillian Dobbie, and Gerald Weber. A lightweight stream-based join with limited resource consumption. In *14th International Conference on Data Warehousing and Knowledge Discovery*, pages 431–442. Springer, 2012, (*ranked B in CORE*).

41. **M. Asif Naeem**, Saif Ullah, and Imran Sarwar Bajwa. Interacting with data warehouse by using a natural language interface. In *17th International Conference on Application of Natural Language to Information Systems*, pages 372–377. Springer, 2012.

42. **M. Asif Naeem**, Gillian Dobbie, Imran Sarwar Bajwa, and Gerald Weber. Resource optimization for processing of stream data in data warehouse environment. In *Proceedings of the International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 62–68. ACM, 2012.

43. **M. Asif Naeem** and Imran Sarwar Bajwa. Generating OLAP queries from natural language specification. In *International Conference on Advances in Computing, Communications and Informatics*, pages 768–773. ACM, 2012.

44. **M. Asif Naeem**, Gillian Dobbie, Gerald Weber, and Imran Sarwar Bajwa. Efficient usage of memory resources in near-real-time data warehousing. In *8th International Multi Topic Conference*, pages 326–337. Springer, 2012.

45. **M. Asif Naeem**, Gillian Dobbie, Gerald Weber, and Imran Sarwar Bajwa. A parametric analysis of stream based joins. In *8th International Multi Topic Conference*, pages 314–325. Springer, 2012.

46. Kashif Hameed, Imran Sarwar Bajwa, and **M. Asif Naeem**. A novel approach for automatic generation of UML class diagrams from XML. In *8th International Multi Topic Conference*, pages 164–175. Springer, 2012.

47. **M. Asif Naeem**, Gillian Dobbie, and Gerald Weber. Optimised X-HYBRIDJOIN for near-real-time data warehousing. In *23rd Australasian Database Conference*, pages 21–30. Australian Computer Society, Inc., 2012, (*ranked B in CORE*).

48. **M. Asif Naeem**, Gillian Dobbie, and Gerald Weber. X-HYBRIDJOIN for near-real-time data warehousing. In *28th British National Conference on Databases*, pages 33–47. Springer, 2011, (*ranked B in CORE*).

49. Imran Sarwar Bajwa and **M. Asif Naeem**. On specifying requirements using a semantically controlled representation. In *16th International Conference on Application of Natural Language to Information Systems*, pages 217–220. Springer, 2011.

50. Ashfa Umber, Imran Sarwar Bajwa, and **M. Asif Naeem**. NL-based automated software requirements elicitation and specification. In *1st International Conference on Advances in Computing and Communications*, pages 30–39. Springer, 2011.

51. Imran Sarwar Bajwa, **M. Asif Naeem**, Ahsan Ali Chaudhri, and Shahzad Ali. A controlled natural language interface to class models. In *13th International Conference on Enterprise Information Systems (ICEIS)*, pages 102–110, 2011, (*ranked C in CORE*).

52. **M. Asif Naeem**, Gillian Dobbie, Gerald Weber, and Shafiq Alam. R-MESHJOIN for near-real-time data warehousing. In *13th international workshop on Data warehousing and OLAP*, pages 53–60. ACM, 2010, (*ranked B in CORE*).

53. Shafiq Alam, Gillian Dobbie, Patricia Riddle, and **M. Asif Naeem**. Particle swarm optimization based hierarchical agglomerative clustering. In *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, volume 2, pages 64–68. IEEE, 2010, (*ranked B in CORE*).

54. Shafiq Alam, Gillian Dobbie, Patricia Riddle, and **M. Asif Naeem**. A swarm intelligence based clustering approach for outlier detection. In *IEEE Congress on Evolutionary Computation (CEC)*, pages 1–7. IEEE, 2010,(*ranked B in CORE*).

55. **M. Asif Naeem**, Gillian Dobbie, and Gerald Weber. Comparing global optimization and default settings of stream-based joins. In *International Workshop on Business Intelligence for the Real-Time Enterprise, (BIRTE, VLDB Workshop)*, pages 155–170. Springer, 2009.

56. **M. Asif Naeem**, Gillian Dobbie, and Gerald Webber. An event-based near real-time data integration architecture. In *12th Enterprise Distributed Object Computing Conference Workshops*, pages 401–404. IEEE, 2008.

57. **M. Asif Naeem**, Imran Sarwar Bajwa, and M. Abbas Choudhary. Hidden web data processing for knowledge management. In *International Conference on Advanced Computer vision and Information Technology*, pages 397–405. IEEE, 2007.

58. Imran Sarwar Bajwa, **M. Asif Naeem**, and Muhammad Nawaz. Web information mining framework using XML based knowledge representation engine. In *International Conference on Software Engineering (ISE 06) Lahore Pakistan, ISE*, volume 6. Citeseer, 2006.

59. Imran S Bajwa, **M. Asif Naeem**, and M A Choudhary Riaz-Ul-Amin. Speech language processing interface for object-oriented application design using a rule-based framework. In *4th International Conference on Computer Applications*, 2006.

Book Chapters

1. **M. Asif Naeem** and Noreen Jamil. Online processing of end-user data in real-time data warehousing. In *Improving Knowledge Discovery through the Integration of Data Mining Techniques*, pages 13–31. IGI Global, 2015.
2. **M. Asif Naeem**, Gillian Dobbie, and Gerald Weber. Big data management in the context of real-time data warehousing. In *Big Data Management, Technologies, and Applications*, pages 150–176. IGI Global, 2014.
3. **M. Asif Naeem**, Gillian Dobbie, and Gerald Weber. Efficient processing of stream data over persistent data. In *Big Data Computing*, page 315. CRC Press, 2013.

Technical Reports

1. **M. Asif Naeem**, Gillian Dobbie, and Gerald Weber. HYBRIDJOIN for near real-time data warehousing. In *Tech. Rep.* University of Auckland, 2010.
2. **M. Asif Naeem**, Gillian Dobbie, and Gerald Weber. X-HYBRIDJOIN for near real-time data warehousing. In *Tech. Rep.* University of Auckland, 2010.

Thesis

1. **M. Asif Naeem**. *Efficient joins to process stream data.* PhD thesis, ResearchSpace@ Auckland, 2012.

Teaching Interests

Data Science, Machine Learning, Deep Learning, Databases, Data Warehousing & Business Intelligence, Data Mining, Big Data, Data Analytics, Data Structure and Algorithms, Object Oriented Programming, Research Methods.

Courses Taught

Teaching at NUCES

2023	**DS3003 – Data Warehousing & Business Intelligence** (Undergrad levl)	88%(positive feedback)
	DS3002 – Data Mining (Undergrad levl)	93%
2022	**CS4033 – Data Warehousing** (Undergrad levl)	95%
	DS3003 – Data Warehousing & Business Intelligence (Undergrad level)	95%
	CS6025 – Advanced Machine Learning (PhD level)	98%
	AI503 – Advanced Machine Learning (Master level)	90%
	CS203 – Database Systems (Undergrad level)	81%
2021	**CS4033(C) – Data Warehousing** (Undergrad level)	93%
	DS3003(G) – Data Warehousing & Business Intelligence (Undergrad level)	83%
	CS203 – Database Systems (Undergrad level)	95%
	AI503 – Advanced Machine Learning (Master level)	72%
2020	**CS203(C) – Database Systems** (Undergrad level)	90%
	CS203(G) – Database Systems (Undergrad level)	85%
	CS408 – Data Warehousing (Undergrad level)	85%
2019	**DS504 – Data Stream Management for Data Science** (Master level)	95%
	CS408(A) – Data Warehousing (Undergrad level)	80%
	CS408(B) – Data Warehousing (Undergrad level)	80%

Teaching at Auckland University of Technology

2018	**COMP810 – Data Warehousing and Big Data** (Postgraduate course)	83%
2017	**COMP810 – Data Warehousing and Big Data** (Postgraduate course)	94%
	INFS601 – Logical Database Design (Undergrad course)	80%
	STAT995– Research Project (Postgraduate course)	93%
2016	**COMP713 – Distributed and Mobile Systems** (Undergrad course)	95%
	COMP810 – Data Warehousing and Big Data (Postgraduate course))	94%
	INFS601 – Logical Database Design (Undergrad course)	93%
2015	**COMP713 – Distributed and Mobile Systems** (Undergrad course)	94%
	COMP810 – Data Warehousing and Big Data (Postgraduate course)	97%
	INFS601 – Logical Database Design (Undergrad course)	92%
2014	**ENSE800 – Software Engineering for Services** (Postgraduate course)	100%
	INFS601 – Logical Database Design (Undergrad course)	93%
2013	**ENSE800 – Software Engineering for Services** (Postgraduate course)	98%
	COMP704 – Research & Development Project (Undergrad course)	95%
	INFS601 – Logical Database Design (Undergrad course)	92%
2012	**INFS500 – Enterprise Systems** (Undergrad course)	91%
	INFS601 – Logical Database Design (Undergrad course)	89%

Teaching at Queens Academic Group

2012	**Q703 – Object Oriented Paradigm** (Level 7 paper)
	Q701 – Logical Databases and Techniques (Level 7 paper)
2011	**Q703 – Object Oriented Paradigm** (Level 7 paper)
	Q604 – Logical Databases and Design (Level 6 paper)
2010	**Q502 – Basic Programming** (Level 5 paper)

Teaching at The University of Auckland

2010	**COMPSCI 351 – Fundamentals of Database Systems** (Undergrad course)
2009	**COMPSCI 105 – Principles of Computer Science** (Undergrad course)
2008	**COMPSCI 101 – Principles of Programming** (Undergrad course)

Plus 7 years Pre-PhD university level teaching experience.

Research Projects

On Going Projects

1. **Digitizing Medical Handwritten Prescriptions**
   The primary aim of this research project is to develop a system prototype for converting the healthcare handwritten notes into digital form using AI and Deep Learning techniques. This digital information will be stored into fully normalized/structured format in order to enrich the existing healthcare data with valuable information. The other aim is to analysis healthcare data from various angles (e.g., patients health patterns) particularly in COVID context using AI techniques. The project is funded by Higher Education Commission of Pakistan (HEC) under his NRPU grant.

2. **Unified VQA**
   Visual Query Answering (VQA) has transformed Information retrieval from visual inputs such as videos and images, but still their capabilities remains limited when applying to documents that include math equations,

ratings and checklists. Such types of documents are very common in industry and education. To the best of our knowledge no single approach has been published in the literature to perform VQA on such documents. In this work we tackled the stated-challenges by developing a unified VQA, an AI-based ensemble approach that combines multiple models to perform VQA on math equations, ratings, and checklists, along with conventional elements of document such as text, table, and figure. The proposed approach utilized SVM for classification of user query, Florance2 to VQA on text, table and figure, DONUT to VQA on ratings and checklists, and MAP to VQA on math equations.

3. **ImageMart**
   ImageMart aims to redefine the e-commerce landscape by introducing an innovative image-based product search system, addressing the limitations of traditional text-based search methods. The significance of ImageMart extends beyond improving the E-Marketplace experience. It introduces a paradigm shift by making online shopping more accessible and engaging for a broader audience. By breaking down language barriers, ImageMart envisions a future where users can effortlessly explore a diverse range of products, furniture and its diverse categories. ImageMart represents a transformative step towards a more inclusive and visually-driven era of online shopping. The project is sponsored by our industry partner, Upstart Commerce.

4. **CaptionCraft**
   The project aim is to tackle the challenge of developing a specialized image captioning model, tailored to retail products such as furniture and mattresses. Unlike conventional image captioning models, our approach does not require pre-written text or annotations from retailers. Instead, retailers simply provide product images, and our model autonomously generates descriptive captions. The significance of this research lies in its pioneering approach, offering a practical and time-saving solution for retailers. By automating the description generation and form-filling processes, we aim to reduce the time and effort required for maintaining product listings. This, in turn, not only increases operational efficiency for retailers but also contributes to a more informative and seamless online shopping experience for customers. The project is sponsored by our industry partner, Upstart Commerce.

5. **VoxAISQL**
   The project focuses on developing a voice-to-SQL (VoxAISQL) interface that simplifies database management by enabling users to interact with databases using natural language. The primary objective of this project is to design and implement a system that converts voice inputs into SQL queries through an AI-powered chatbot. This system will allow users to query databases and retrieve data without needing to understand SQL, providing an accessible and efficient way to manage data. The project aims to empower non-technical users, businesses, and individuals with disabilities by making database management more intuitive and accessible.

6. **Your AI Assistant (YAIA)**
   YAIA is a commanding application utilizing advanced AI technologies. Its purpose is to take in user queries in the form of speech as well as text. YAIA then provides output in the form of text, speech, or some desired action. YAIA excels in user-friendliness, making it accessible and easy to use for a broad audience. Its interface is designed to be intuitive, allowing both tech-savvy individuals and those less familiar with technology to navigate effortlessly. With options for text or voice commands, it ensures accessibility for users regardless of their preferred mode of interaction. The voice-activated features not only enhance accessibility for those with physical limitations but also cater to users seeking hands-free operation, promoting convenience and multitasking.

7. **DiagnosysAI**
   DiagnosysAI goes beyond traditional healthcare solutions, offering a diagnostic system with unparalleled benefits for individuals and the community alike. Imagine a healthcare experience where you, as an individual, have the power to input your symptoms in a simple and user-friendly interface. The system, equipped with advanced algorithms, intelligently dissects these symptoms to identify the presence of various diseases. This means personalized recommendations for separate tests, ensuring a precise and efficient diagnostic journey tailored to your unique health profile. The layman gains access to comprehensive disease insights, empowering them with knowledge and facilitating more informed discussions with healthcare professionals. Beyond individual benefits, the system contributes to the community by optimizing healthcare resources. With reduced unnecessary testing and streamlined diagnostics, it paves the way for a more efficient and accessible healthcare system, positively impacting the well-being of the entire community. DiagnosysAI represents an important step towards personalized, community-centric healthcare, revolutionizing the way we approach and experience medical diagnostics.

8. **Bil-FaND: Multi Features based Bi-Lingual Fake News Detection**
   The project aim is to detect bilingual fake news using multiple features such as textual, numeric, categorical, and multimedia. To achieve this, we used multiple models e.g. LSTM for numeric and categorical, multilingual BERT for both English and Urdu text, and BLEU for the multimedia. The outcomes of the numerical, textual, category, and caption-generating layers improved the accuracy and reliability of detecting fake news. This study has practical implications for countering misinformation in multilingual contexts and advances natural language processing approaches in the context of fake news identification. The project is sponsored by Government of Pakistan.

9. **Transformation Accelerator**
   In Real-Time Data Warehousing (RTDM) typically a fast input stream of customer's sales transactions from Operational Data Sources (ODSs) needs to be joined with a disk-based Master Data (MD) before loading it to RTDM. These semi-stream join approaches typically perform the join with a limited main memory partition assigned to them, which is generally not large enough to hold the whole MD. In this project we present novel approaches for caching and load shedding in semi-stream one-to-many equijoins. The project is implemented in Mitre10, one of the biggest hardware warehouse chain in New Zealand.

10. **Personalized Taste Profiling**
    The aim of this research project is in the formulation of taste profiles for microblog users based on the content they disseminate as well as the social network structure. In particular, the research focuses on semantic understanding and relatedness of short, sparse and noisy texts which are challenging for existing algorithms to learn. The project is broadly categorized into two parts (a) extraction of personalized user interests in microblog texts (b) audience measurement (taste profiling) based on the temporal-dynamism of the social network structure in microblogs. Twitter will provide the testing platform for our approaches mainly because of its popularity, API access ability as well as the dynamism of its overall network structure. This research is fundamental to services and content recommendations and audience measurement e.g. to marketers.

Completed Projects

2023	ViLanOCR: A Multilingual OCR for English and Low Resource Urdu Language
2023	MEST : Multilingual Extraction of Semi-Structured Printed and Handwritten Text
2022	ClickFix: Fixing your Home Appliances at One Click
2021	A Deep Level Basket Analysis
2021	Smart Shopping
2020	Parallelising Stream-Relation Join for Near-Real-Time Data Warehouse
2019	Predicting User Personality from Public Perceptions on Social Media
2019	Extracting Data from Digital and Scanned Charts
2018	Monitoring Health using IoT
2018	Short-term Price Prediction for Customers
2017	Reusing Past Replies to Respond to New Emails
2017	Staged Online Learner (SOL)
2017	Post Discharge Application

Research Talks

1. **Trends and Challenges in Data Science**, Invited talk at NUCES, Islamabad, Pakistan 2020.

2. **Recent Developments in Data Science - in Big Data Perspective**, Invited talk at UAEU, Al Ain, Dubai, 2019.

3. **Updates on Data Science Research Centre (DSRC) activities**, Data Science Research Centre Workshop, Auckland, New Zealand, 2018.

4. **Aspect of Blame in Tweets: A Deep Recurrent Neural Network Approach**, The 2017 World Wide Web conference, Perth, Australia, 2017.

5. **Document Level Semantic Comprehension of Noisy Text Streams via Convolutional Neural Networks**, The IEEE 21st International Conference on Computer Supported Cooperative Work in Design (CSCWD), Wellington, New Zealand, 2017.

6. **Uncovering Useful Patterns in Shopping Cart Data**, The IEEE 21st International Conference on Computer Supported Cooperative Work in Design (CSCWD), Wellington, New Zealand, 2017.

7. **A Review on IoT Healthcare Monitoring Applications and a Vision for Transforming Sensor Data into Real-time Clinical Feedback**, The IEEE 21st International Conference on Computer Supported Cooperative Work in Design (CSCWD), Wellington, New Zealand, 2017.

8. **Recent Developments in Data Science**, Keynote Speech at The 6th International conference on Innovative Computing Technology (INTECH 2016), Islamabad, Pakistan, 2016.

9. **Optimising Queue-Based Semi-stream Joins by Introducing a Queue of Frequent Pages**, The 27th Australasian Database Conference (ADC), Sydney, Australia, 2016

10. **Research Challenges in Processing of Streaming Data particularly in the context of RDW**, Data Science Research Group Workshop, Auckland, New Zealand, 2016.

11. **Caching and Load Shedding in Semi-Stream Joins for Skewed Big Data**, eResearch NZ, Queenstown, 2016.

12. **Data Science Research Centre (DSRC) Launch**, Auckland, New Zealand, 2015.

13. **A Cache-based Semi-Stream Join to deal with Unmatched Stream Data**, The 26th Australasian Database Conference, ADC 2015, Melbourne, Australia 2015.

14. **S3J: A Parallel Semi-Stream Similarity Join**, 18th ACM International Workshop On Data Warehousing and OLAP, Melbourne, Australia, 2015.

15. **Memory Efficient Join to Process Semi-Stream Data**, Centre for Artificial Intelligence Research (CAIR), AUT, 2015.

16. **Efficient Processing of Semi-Stream Data**, Knowledge Engineering and Discovery Research (KEDRI), AUT, 2015.

17. **Optimizing Queue-based Semi-Stream Joins with Indexed Master Data**, 16th International Conference on Data Warehousing and Knowledge Discovery, Munich, Germany, 2014.

18. **A Caching Approach to Process Stream Data in Data Warehouse**, The 9th International Conference on Digital Information Management, Bangkok, Thailand, 2014.

19. **Natural Language Processing Based Interpretation of Skewed Graphs**, The 3rd International Conference on Advances in Computing, Communications and Informatics, Delhi, India, 2014.

20. **A Generic Front-Stage for Semi-Stream Processing**, The 22st ACM International Conference on Information and Knowledge Management (CIKM), San Francisco, CA, USA, 2013.

21. **SSCJ: A Semi-Stream Cache Join using a Front-Stage Cache Module**, 15th International Conference on Data Warehousing and Knowledge Discovery (DaWaK), Prague, Czech Republic, 2013.

22. **Efficient Joins to Process Semi-Stream Data**, Keynote Speech, The 9th International Conference on Digital Information Management, Islamabad, Pakistan, 2013.

23. **Tuned X-HYBRIDJOIN for Near-Real-Time Data Warehousing**, The 15th Asia-Pacific Web Conference (APWeb), Sydney, Australia, 2013.

24. **Optimised X-HYBRIDJOIN for Near-Real-Time Data Warehousing**, The 23rd Australasian Database Conference (ADC), Melbourne, Australia, 2012.

25. **A Lightweight Stream-based Join with Limited Resource Consumption**, 14th International Conference on Data Warehousing and Knowledge Discovery, Vienna, Austria, 2012.

26. **Stream-based Joins with Limited Resource Consumption**, Database Systems and Information Management (DIMA) Group, TU Berlin, 2012.

27. **A Parametric Analysis of Stream Based Joins**, 8th International Multi-topic Conference, Pakistan, 2012.

28. **Efficient Usage of Memory Resources in Near-Real-Time Data Warehousing**, 8th International Multi-topic Conference, Pakistan, 2012.

29. **X-HYBRIDJOIN for Near-real-time Data Warehousing**, 28th British National Conference on Databases (BNCOD), Manchester, 2011.

30. **R-MESHJOIN for Near-Real Time Data Warehousing**, 13th International Workshop On Data Warehousing and OLAP (DOLAP), Toronto, Canada, 2010.

Supervision & Mentoring Experience

Current Students

PhD

1. **Waheed Ashraf**, Primary Supervisor, Topic of research: *High Utility and Frequent Itemset Mining*

2. **Farhana Sharief**, Co-Supervisor, Topic of research: *Multi-class Imbalanced Data Handling in Fog Computing*

3. **Widana Kankanamge Darsha Jayamini**, Co-Supervisor, Topic of research: *Triangulating an Explainable Personalized Asthma Risk Control (PARC) Score using Multiple Data Sources*

4. **Rashi Bhalla**, Co-Supervisor, Topic of research: *Novel Methods for Distributed Data Stream Mining*

Master

1. **Warsha Khan**, Primary Supervisor, Topic of research: *Query-based Text Extraction from Medical Handwritten Documents*

2. **Saad Munir**, Primary Supervisor, Topic of research: *BiL-FaND: An Ensembled Bi-Lingual Fake News Detection Approach*

3. **Abdul Muqtadir**, Primary Supervisor, Topic of research: *An AI-based Approach for Query Answering (Q/A) on Unstructured Multi-Layout Printed Scanned Documents*

4. **Humamayoun Mustafa Mazhar**, Primary Supervisor, Topic of research: *Disambiguate Medicine Names in Handwritten Medical Prescription using Deep Learning*

5. **Muhammad Aamir Gulzar**, Primary Supervisor, Topic of research: *A Light Weight Image Processing Based Technique to Extract Data from Line Charts*

6. **Shaina Laraib**, Primary Supervisor, Topic of research: *Key-Word Spotting in Low Resource Speech*

Past Students (selected ones)

1. **Wandabwa Masindano Herman**, Primary Supervisor, Topic of research: *Deep learning Methods in Understanding Noisy Data Streams.*

2. **Benjamin James Denham**, Co-supervisor, Topic of research: *Novel Methods for Case Classification from Natural Language Descriptions.*

3. **Alhanof AlMutairi**, Co-supervisor, Topic of research: *Towards Enterprise Systems Adaptability - effects of software-technological choices on cost and reliability risks in business process management.*

4. **Waruni Hewage**, Co-supervisor, Topic of research: *Trading of Privacy Against Accuracy in a Data Stream Environment.*

5. **Jehan Badshah**, Co-supervisor, Topic of research: *Load Optimization in Smart Grid System by Evolutionary Computing.*

6. **Hoa Hong Nguyen**, Co-supervisor, Topic of research: *An intelligent falls management system: A field study.*

7. **Nawal Chanane**, Co-supervisor, Topic of research: *The Influence of Mobile Apps & Social Networking in Promoting Medication Adherence.*

8. **Komala Dhanapal Sagadevan**, Co-supervisor, Topic of research: *Predicting Stock Market Index Using Artificial Intelligence Techniques.*

9. **Chamari Indunil Kithulgoda**, Co-supervisor, Topic of research: *Staged Online Learning Approach for Concept Drifting and Recurring Data Streams.*

10. **Hira Tahir**, Primary Supervisor, Topic of research: *An Automatic Textual Highlights Generation for Episode-wise Urdu Dramas using YouTube Meta Data.*

11. **Umair Ahmad**, Primary Supervisor, Topic of research: *EMT-DocNet: A Framework for Extracting Marked Data Objects from Unstructured Visually Rich Documents.*

12. **Huy Vuong Nguyen**, Primary Supervisor, Topic of research: *Predicting of right time to buy a product.*

13. **Wayan Linggawa**, Primary Supervisor, Topic of research: *Reusing Past Replies to Respond to New Emails: A Case-based Reasoning Approach.*

14. **Zheng Ma**, Primary Supervisor, Topic of research: *Forensic Social Media Analytics.*

15. **Eliazar Dimalapang**, Primary Supervisor, Topic of research: *Forensic Intelligence.*

16. **Kim Tung Nguyen**, Primary Supervisor, Topic of research: *Multi-way Semi-stream Join for Near-real-time Data Warehousing.*

17. **Lijuan Ning**, Primary Supervisor, Topic of research: *The Five-minute Rule 30 Years Later and the Impact of SSD on database.*

18. **Victoria Ponomareva**, Secondary Supervisor, Topic of research: *Script Based Feature Identification for TV Series Uptake Prediction.*

19. **Junshu Wang**, Secondary Supervisor, Topic of research: *Development of a Model to Predict TV Serial Uptake based on Pre-production Information.*

20. **Erum Mehmood**, Primary Supervisor, Topic of research: *Optimization of Cache-based Semi-Stream Join.*

21. **Omer Aziz**, Primary Supervisor, Topic of research: *Optimized Approach for Accessing of Disk-Based Master Data in HYBRIDJOIN (Hybrid Join).*

22. **Solomon George**, Primary Supervisor, Topic of research: *Performance Evaluation and Extension of Cache Join in Real-life Environment.*

23. **Kingsley Basker**, Primary Supervisor, Topic of research: *Performance Investigation of Stream-based Joins using Different Storage Media.*

24. **Linda Wang**, Secondary Supervisor, Topic of research*Frequent Item-set Mining with Master Data.*

25. **Xiao Guo**, Secondary Supervisor, Topic of research:*Cost Based Stream-Relation Join Algorithm Selection.*

26. **Muhammad Anees**, Primary Supervisor, Topic of research: *Implementing Cache Efficiently in Stream-Realtion Join.*

27. **Ashar Nawaz**, Primary Supervisor, Topic of research: *Building an Ontology using Crimes Related Social Data to Help in Finding Similar Crimes.*

28. **Wasiullah Waqar**, Primary Supervisor, Topic of research: *Deep Level Market Basket Analysis.*

29. **Adi Darliansyah**, Primary Supervisor, Topic of research: *Incorporating User Personality in Public Perceptions on Twitter.*

30. **Yiwei Feng**, Primary Supervisor, Topic of research: *A Smart Email Reply Client using Deep Learning Approach.*

31. **Kathleen Silva de Azevedo**, Secondary Supervisor, Topic of research: *Applying artificial neural networks to chart recognition.*

Professional Activities

Project Reviewer

1. Project reviewer for Knowledge Foundation Center, Sweden.
2. Project reviewer for Punjab Higher Education Commission.
3. Project reviewer for Sino-Pak Center for Artificial Intelligence (SPCAI).
4. Self Assessment Reviewer, Department of Data Science, NUCES

Members in Committees

1. Member of MSRC Committee, Department of AI&DS, NUCES
2. Member of Discipline Committee, Department of Computer Science, NUCES
3. Master & FYP evaluation committee, Department of Computer Science, NUCES
4. Member of Board of Study, School of Engineering, Computer and Mathematical Sciences
5. Member of Event Organising Committee, at The University of Auckland

Keynote Speaker

1. **Title:** *Recent Developments in Data Science*, in INTECH 2016.
2. **Title:** *Efficient Joins to Process Semi-Stream Data*, in ICDIM 2013.

Organising Role

1. **Conference Chair**, INMIC 2022.
2. **Associate Editor**, IEEE Access Journal.
3. **Guest Editor**, Special Issue: Real-time Data Management and Analytics in Electronic Journal.
4. **Track Chair**, Conf-IRM 2019.
5. **General Co-chair**, INTAP 2019.
6. **Publicity Chair**, UCC/BDCAT 2019.
7. **Proceedings Editor**, ICDIM 2017 and 2018.
8. **Contest Chair**, PAKDD 2016.
9. **Organising Chair**, IEEE-IWDS since 2013.

10. **Workshop Organiser**, Data Science Research Centre (DSRC) since 2013.
11. **Workshop Chair**, ICDIM 2015.
12. **General Track Chair**, ICDIM 2013.

Review Member in Journal and Conferences

1. Expert Systems with Applications.
2. Data & Knowledge Engineering.
3. IEEE Transaction on Very Large Scale Integration (VLSI) Systems.
4. Neurocomputing Journal.
5. Computers and Electrical Engineering Journal.
6. Journal of Software: Evolution and Process.
7. 31st Australasian Database Conference (ADC) (2015 and onward).
8. 22nd IEEE International Multi Topic Conference (INMIC 2019).
9. 51th Hawaii International Conference on System Sciences (HICSS 2017).
10. ICIS 2017.
11. AMCIS 2017.
12. 9th Asian Conference on Intelligent Information and Database Systems.
13. The Australasian Computer Science Conference (ACSC), (2013 and onward).
14. First New Zealand Text Mining Workshop (2016).
15. IEEE International Conference on Digital Information Management (2016), (2013 and onward).

Development Experience

2019	Project title: *A Smart price predictor*, Published in C&EE.
2019	Project title: *SENTIPEDE: A Smart System for Sentiment-based Personality Detection from Short Texts*, Published in JUCS.
2018	Project title: *A Smart Email Client*, Published in IEEEAccess.
2013	Project title: *A generic Front-Stage for semi-stream joins processing*, Published in CIKM 2013.
2012	Project title: *CACHEJOIN for generalized data stream processing*, Published in DaWaK 2012.
2011	Project title: *X-HYBRIDJOIN for skewed data stream processing*, Published in BNCOD 2011.
2010	Project title: *R-MESHJOIN for real-time ETL*, Published in DOLAP 2010.

Programming Skills

LAVA2, LLaMA, GPT3.5, DONUT, Pix2Struct, LayoutLM, DESSURT, BROS, TrOCR, PPOCR, NER, LSTM, VADER, BERT, LDA, DT, SVM, VADER, CNN, RCNN, LSTM, GRU, and GLOVE. Java, Python, Apache Hadoop, Apache Flink.

References

References will be provided on demand.

1. Dr. Gillian Dobbie, Professor, School of Computer Science, The University of Auckland, New Zealand
email: g.dobbie@auckland.ac.nz.

2. Dr. Dave Parry, Professor, School of Information Technology, Murdoch University, Australia
email: David.Parry@murdoch.edu.au.

3. Dr. Gerald Weber, Senior Lecturer, School of Computer Science, The University of Auckland,
email: g.weber@auckland.ac.nz