



# Souvik Kundu

## Curriculum Vitae

### Areas of Interest

Deep Learning, Natural Language Processing, Natural Language Understanding, Large Language Models, Deep Reinforcement Learning

### Education

- 2014–2020 **PhD in Computer Science**, *National University of Singapore*, Singapore, CGPA - 4.75/5.0  
**Thesis Title** - Question Answering Using Deep Neural Networks: Single Turn and Beyond
- 2010–2014 **Bachelor of Engineering in Electronics and Telecommunication**, *Jadavpur University*, Kolkata, India, CGPA - 9.06/10.0

### Experience

- 2020–Present **Senior Applied Scientist**, *Microsoft*, Redmond, USA
- 2023 **Enhancing LLMs with Dynamic Few-shots**, *Developed a target similarity tuning model to dynamically select relevant few-shot examples, enabling the creation of tailored prompts for GPT and ChatGPT-based NL2SQL models. This resulted in significant improvements ( $> 5\%$  absolute) in accuracy over traditional static few-shot setup.*
- 2022 **Retriever Augmented LLMs**, *Actively working in research, design and implementation of retriever augmented Large Language Models (LLMs) to optimize information retrieval, refine context understanding, and achieve superior performance in complex natural language processing tasks.*
- 2022 **Schema Matching/Form Key mapping**, *Conducted extensive research and developed a production-quality schema matching solution using advanced Transformer models from scratch and successfully passed public preview (an active feature in today's Microsoft Form Recognizer product).*

Microsoft – Redmond – WA, USA

✉ [sk210892@gmail.com](mailto:sk210892@gmail.com) • [in](https://www.linkedin.com/in/souvik-kundu-ph-d-04bab493/) [souvik-kundu-ph-d-04bab493/](https://www.linkedin.com/in/souvik-kundu-ph-d-04bab493/)  
🔗 [sk210892](https://scholar.google.com/citations?user=sk210892) • [Google Scholar](https://scholar.google.com/citations?user=sk210892)

- 2021 **Model-based Schema Linking**, Developed a siamese-style schema linking model utilizing pre-trained sentence embedding models to identify top-k relevant columns for converting natural language questions into structured queries. This contribution empowered the product to effectively handle complex queries, while achieving a remarkable reduction of over 25% in the end-to-end system latency.
- 2020 **Hybrid Ranking Network for NL2SQL**, Contributed to the development of a Hybrid Ranking Network (HydraNet) for NL2SQL tasks, involving column-wise ranking and decoding to assemble column outputs into SQL queries. Published in Arxiv and achieved top performance on the WikiSQL leaderboard for a significant period.

- 
- 2018 **Research Intern**, Allen Institute for Artificial Intelligence (AI2), Seattle, USA
- Multi-hop Question Answering**, Proposed a novel approach for multi-hop QA that involves generating potential paths, extracting implicit relations along these paths, and composing them to encode each path. This contribution led to a publication in ACL 2019, showcasing the innovative research conducted.

- 
- 2014-2020 **Ph.D. Researcher**, National University of Singapore, Singapore
- 2019 **Conversational Question Answering**, In this work, a new follow-up question identification task was introduced. Also, a three-way attentive pooling network was proposed to capture topic continuity and topic shift while scoring a particular candidate follow-up question. Published in ACL 2020.
- 2018 **Nil-Aware Machine Reading Comprehension**, Proposed an unified framework that can be integrated to many machine comprehension systems. Published in EMNLP 2018.
- 2017 **Reading Comprehension-based Question Answering**, Proposed a neural model to tackle the complex co-reference resolution and for implicit understanding of answer type. Published in AAAI 2018.
- 2016 **Noise Robust DNN-based ASR**, To improve the adaptability and interpretability of DNN-based acoustic models, a Vector Tailor Series (VTS) compensated Gaussian Mixture Model is combined with a DNN. Published in INTERSPEECH 2016.
- 2015 **Joint Acoustic Factor Learning for Robust ASR**, Completed this project in the Jelinek Summer workshop of Speech and Language Technology (JSALT 2015). Explored the use of discriminative auxiliary input features obtained using joint acoustic factor learning for Deep Neural Network adaptation. Published in ICASSP 2016.

---

## Awards and Achievements

- 2018 **Research Achievement Award**, National University of Singapore
- 2014 - 2018 **Research Fellowship**, National University of Singapore

Microsoft – Redmond – WA, USA

✉ sk210892@gmail.com • in souvik-kundu-ph-d-04bab493/  
🔗 sk210892 • Google Scholar

---

## Selected Publications

- [1] Qin Lyu, Kaushik Chakrabarti, Shobhit Hathi, Souvik Kundu, Jianwen Zhang, and Zheng Chen. Hybrid ranking network for text-to-sql, <https://arxiv.org/abs/2008.04759>, 2020.
- [2] Qian Lin, Souvik Kundu, and Hwee Tou Ng. A co-attentive cross-lingual neural model for dialogue breakdown detection, Proceedings of the 28th International Conference on Computational Linguistics (COLING), 2020.
- [3] Souvik Kundu, Qian Lin, and Hwee Tou Ng. Learning to identify follow-up questions in conversational question answering, Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL), 2020.
- [4] Souvik Kundu, Tushar Khot, Ashish Sabharwal, and Peter Clark. Exploiting explicit paths for multi-hop reading comprehension, Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL), 2019.
- [5] Souvik Kundu and Hwee Tou Ng. A nil-aware answer extraction framework for question answering, Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2018.
- [6] Souvik Kundu and Hwee Tou Ng. A question-focused multi-factor attention network for question answering, Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, 2018.
- [7] Souvik Kundu, Khe Chai Sim, and Mark Gales. Incorporating a generative front-end layer to deep neural network for noise robust automatic speech recognition, 17th Annual Conference of the International Speech Communication Association (INTERSPEECH), 2016.
- [8] Souvik Kundu, Gautam Mantena, Yanmin Qian, Tian Tan, Marc Delcroix, and Khe Chai Sim. Joint acoustic factor learning for robust deep neural network based automatic speech recognition, 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016.

---

## Professional Academic Services

- 2018-2023 **Computational Linguistics (CL)**, *Reviewer*
- 2021 **North American Association of Computational Linguistics (NAACL)**, *Program Committee Member*
- 2018-2020 **Association of Computational Linguistics (ACL)**, *Program Committee Member*
- 2018 **Journal of Natural Language Engineering (JNLE)**, *Reviewer*

---

## Technical Skills

- DL Tools    PyTorch, Tensorflow, Hugging-face, AllenNLP, Keras    Language    Python, C#
- Infrastructure    Azure ML, AWS Fullstack (Sagemaker, API Gateway, IAM, S3, etc.)

Microsoft – Redmond – WA, USA

✉ [sk210892@gmail.com](mailto:sk210892@gmail.com) • **in** [souvik-kundu-ph-d-04bab493/](https://www.linkedin.com/in/souvik-kundu-ph-d-04bab493/)  
🔗 [sk210892](https://scholar.google.com/citations?user=sk210892) • [Google Scholar](https://scholar.google.com/citations?user=sk210892)