Sanket Shah

Senior Research Scientist
Salesken AI
Bangalore, KA, India.

✉ ssanket369@gmail.com
⌂ https://sanket0211.github.io/
🎓 Google Scholar

EDUCATION

- **International Institute of Information Technology - Hyderabad** *2014 - 2019*
  B.Tech and MS by Research in Computer Science CGPA: 9/10
  Advisor: Prof. Nita Parekh

RESEARCH EXPERIENCE

- **Saarthi AI - *Senior Research Scientist*** *July 2022 - Present*
  Currently, I am working on a couple of project at Saarthi AI. One of them is developing continous streaming Automatic Speech Recognition (ASR) system for Indic Languages. This includes:
  - Preparing data pipe-line for storing, cleaning & annotating speech data.
  - Customizing Microsoft Azure Speech-to-text (STT) system using text/Audio data. Experiments include estimating interpolation weight between the generic LM and Biased LM for boosting speech recognition for in-domain data.
  - Experimenting with word-level confidence scores for improving precision of the recognized words.
  - Integrating Facebook Denoiser with streaming STT for removing background noise like traffic, cross-talks, fan, etc and improving speech recognition.
  - Also developing Kaldi based speech recognition from scratch using in-house data and comparing its performance with that of custom Azure STT model. I am also working on building in-domain dialogue system where we predict BOT's response to user utterance. - Implemented various statistic methods like decision trees, random forest, XG Boost and Naive Bayes and compared with neural network and deep learning based architectures like BERT/Transformers for classification.

- **Salesken AI - *Senior Research Scientist*** *July 2021 - June 2021*
  I work on a number of NLP downstream tasks like text-to-text translation, semantic similarity, emotion recognition, sentiment analysis, and text summarization. This includes training ML models from scratch, designing evaluation strategies and optimizing & re-structuring the code for final deployment.
  - Developed model for semantic similarity using sentence-transformers from Hugging Face.
  - Built a sequence-to-sequence model for understanding conversation context and predicting spoken dialogues.
  - Created customized versions of massive DL models by simplifying network architectures and reducing model size using techniques like ONNX, TensorRT and model quantization.
  - Involved in end-to-end model deployment involving API creation, decker images, and creating deployment files.

- **Microsoft Research India - *Research Fellow (Speech and NLP Group)*** *July 2019 - June 2021*
  Advisors: Dr. Sunayana Sitaram
  Worked on the problem of Automatic Speech Recognition (ASR) for multilingual and code-switched conversational speech.
  - Improving generalization capability of pre-trained models in new domain data using regularization techniques.
  - Used adversarial training to remove bias during multi-task learning.

- **Microsoft Research India - *Research Intern (Speech and NLP Group)*** *Jan 2019 - June 2019*
  Advisors: Dr. Sunayana Sitaram
  - Developed a system to detect spoken terms in a conversation by simple pre-processing and fuzzy matching.
  - Developed a tool using Django for easing the process of transcribing code-switched speech. Providing ordered list spoken term suggestions to the annotator using speech recognition system.

- **Indian Institute of Science, Bangalore - *Research Intern at MALL Lab, IISc*** *May 2018 - September 2018*
  Advisors: Prof. Partha Pratim Talukdar
  - Worked on problem of entity extraction and linking in knowledge graphs.
  - Built a text based question answering system.

PUBLICATIONS

Preprints * - equal contribution

3. **Curriculum Learning for Adapting Models on Code-Mixed Data in Emotion Recognition**
   Dheeraj Agrawal, **Sanket Sanjay Shah**, Neeru Dubey, Shubham Sharma, Bharatram Natarajan, Suvro Banerjee and Ashish Kumar
   Under Review, COLING 2022

2. **Learning to Recognize Code-switched Speech Without Forgetting Monolingual Speech Recognition.**
Sanket Shah, Basil Abraham, Gurunath Reddy M, Sunayana Sitaram, Vikas Joshi
arXiv:2007.02519, 2020.

1. **Cross-lingual and Multilingual Spoken Term Detection for Low-Resource Indian Languages.**
Sanket Shah, Satarupa Guha, Simran Khanuja, Sunayana Sitaram
arXiv:2007.02519, 2020.

Conference Publications

7. **Multilingual and code-switching ASR challenges for low resource Indian languages**
Anuj Diwan, Rakesh Vaideeswaran, **Sanket Shah**, Ankita Singh, Srinivasa Raghavan, Shreya Khare, Vinit Unni, Saurabh Vyas, Akash Rajpuria, Chiranjeevi Yarra, Ashish Mittal, Prasanta Kumar Ghosh, Preethi Jyothi, Kalika Bali, Vivek Seshadri, Sunayana Sitaram, Samarth Bharadwaj, Jai Nanavati, Raoul Nanavati, Karthik Sankaranarayanan, Tejaswi Seeram, Basil Abraham
Interspeech 2021.

6. **Learning not to Discriminate: Task Agnostic Learning for Improving Monolingual and Code-switched Speech Recognition**
Gurunath Reddy, **Sanket Shah**, Basil Abraham, Vikas Joshi, Sunayana Sitaram.
*Workshop on Speech Technologies for Code-switching in Multilingual Communities (**WSTCSMC**), 2020*

5. **First Workshop on Speech Processing for Code-switching in Multilingual Communities: Shared Task on Code-switched Spoken Language Identification**.
Sanket Shah, Sunayana Sitaram, Rupeshkumar Mehta.
*First Workshop on Speech Processing for Code-switching in Multilingual Communities (**WSTCSMC**), 2020.*

4. **Unsung Challenges of Building and Deploying Language Technologies for Low Resource Language Communities**.
Pratik Joshi, Christain Barnes, Sebastin Santy, Simran Khanuja, **Sanket Shah**, Anirudh Srinivasan, Satwik Bhattamishra, Sunayana Sitaram, Monojit Choudhury, Kalika Bali.
*International Conference on Natural Language Processing (**ICON**), 2019.*

3. **Using monolingual speech recognition for spoken term detection in code-switched hindi-english speech**.
Sanket Shah, Sunayana Sitaram
*International Conference on Data Mining Workshops (**ICDMW**), 2019 .*

2. **CoSSAT: Code-Switched Speech Annotation Tool**.
Sanket Shah, Pratik Joshi, Sebastin Santy, Sunayana Sitaram
*Workshop on Aggregating and Analysing Crowdsourced Annotations for NLP(**EMNLP**), 2019* (**Long Oral**).

1. **Kvqa: Knowledge-aware visual question answering**.
Sanket Shah, Anand Mishra, Naganand Yadati, Partha Pratim Talukdar.
*Association for the Advancement of Artificial Intelligence (**AAAI**), 2019.*

SELECTED RESEARCH PROJECTS

Curriculum Learning for Adapting Models on Code-Mixed Data in Emotion Recognition   *Mar'22 - Present*
Salesken AI

- Proposed curriculum learning for fine-tuning models on code-switched data.
- Proved the sensitivity of multilingual models on script of the language and showed that simple script change of code-switched text can lead to boost in performance.

Learning Without Forgetting for Code-switched Speech Recognition   *Feb 2020 - June 2021*
Advisor: Dr. Sunayana Sitaram, *Microsoft Research India*

- Proposed 'Learning Without Forgetting' technique for code-switched speech recognition system.
- Effectively adapting large speech recognition systems to perform well on code-switched speech.

Learning not to Discriminate - Task Agnostic Learning for monolingual and Code-switched Speech Recognition   *Feb 2020 - June 2020*
Advisor: Dr. Sunayana Sitaram *Microsoft Research India*

- Proposed an adversarial multi-task framework for monolingual and code-switched speech recognition.
- Showed that it is important to train shared parameters which are task agnostic unbiased performance on all the tasks during multi-task learning.

Shared task on Code-switched Spoken Language Identification at the First Workshop on Speech Processing for Code-switching *Feb 2020 - Aug 2020*

Advisor: Dr. Sunayana Sitaram *Microsoft Research India*

- Member of the shared task organizing committee at the First Workshop on Speech Processing for Code-switching.
- Involved in dataset creation, baseline generation and evaluating the models submitted by the participating teams.

KVQA: Knowledge-aware Visual Question Answering *April 2018 - April 2019*

Advisor: Prof. Partha Pratim Talukdar *Microsoft Research India*

- Proposed the largest dataset for Visual Question Answering (VQA) over KG (as on November 14, 2018).
- Proposed solution for an important but unexplored problem of VQA involving named entities in an image.
- Visual entity linking problem in web-scale.
- Challenges for Computer Vision: face identification at web-scale.
- Challenges for NLP: reasoning over KG.

Selected Awards and Honors

- IIIT Hyderabad & Microsoft Research Travel Grant to attend **AAAI '19** *2019*

Professional Responsibilities

- *Undergraduate Teaching Assistantship* - Computer Science and Engineering, IIIT Hyderabad
    - Computer Programming - *Prof. Vineet Gandhi., Prof. Praveen Parchuri* *Fall 2017*
    - Data Structures and Algorithms - *Prof. Lini Thomas* *Spring 2016*

Courses Taken

- *Core Technical Courses:* Computer Programming, Data Structures & Algorithms, Artificial Intelligence, Natural Language Processing, Statistical Methods in AI.
- *Core Systems Courses:* Database Systems, Operating Systems, Optimization Methods, System Design & Engineering.

Technical skills

- **Programming Language:** c, c++, Python
- **Database Systems:** MySQL, PostgreSQL
- **Web Technologies:** HTML, CSS, PHP, JavaScript, jQuery, Django
- **Machine Learning / Data Science Tools:** NumPy, Pandas, Scikit-Learn, Matlab, Tensorflow, PyTorch