Revanth Gangi Reddy

Graduate Research Assistant, UIUC ***Siebel Scholar (Class of 2022)***

✉ g.revanthreddy111@gmail.com • gangiswag.github.io

Interests: Natural Language Processing, Deep Learning and Machine Learning
Citizenship: Indian

Education

B. Tech, Computer Science	Indian Institute of Technology (IIT), Madras	CGPA - 9.16/10.0	July'14 - May'18
MS in Computer Science	University of Illinois at Urbana-Champaign	GPA - 3.89/4.0	Jan'21 - May'22

Work Experience

University of Illinois at Urbana-Champaign **Oct 2020 - Present**
Research Assistant, Blender Lab *Champaign, United States*

- Working on projects in the areas of multimodal question answering, information retrieval and claim detection in news articles. Ongoing collaborations with Amazon, IBM Research, Columbia University and UNC Chapel Hill.

IBM Research AI **Oct 2019 - Oct 2020**
AI Resident, Multilingual NLP team *New York, United States*

- Worked on projects in the areas of question answering, open-domain knowledge retrieval and AMR parsing. Contributed to five research papers (four first-author) and two patents during the AI Residency program.

Microsoft **Oct 2018 - Sep 2019**
Software Engineer, Data Integration team *Vancouver, Canada*

- Worked in the team that develops data connectors for PowerApps, LogicApps and Flow. Built the entire data pipeline for a scalable Azure-based multi-region logging infrastructure that currently handles 7 billion logs a day.

Internships

IBM Research AI **Summer 2018**
Research Intern, Watson Conversations team *New Delhi, India*

- Worked on better neural memory architectures for improving the performance of task-oriented dialog systems. My work during the internship was published as a long paper at NAACL 2019.

Microsoft India Development Center **Summer 2017**
Research Engineering Intern, Cortana Personalization Team *Hyderabad, India*

- Analyzed user behaviour patterns based on user temporal data and developed a model for forecasting user activity given the past history.

MyAlly.ai **Summer 2016**
Summer Intern *Hyderabad, India*

- Worked on named entity recognition for an automated customer meeting scheduler and built the backend for an email interface, similar to gmail, that can support accounts from multiple domains.

Current Research Projects

A New Benchmark for Claim Detection with Background Knowledge
Under submission to ACL Rolling Review

- Proposed an evaluation benchmark that **extends claim detection** to more background attributes. Introduced a zero-shot QA based framework for the task and compared against different few-shot and prompt-based approaches.

Entity-Conditioned Question Generation for Robust Attention Distribution in Neural IR
Under submission to ACL Rolling Review

- Proposed an **entity-conditioned data augmentation** strategy that generates questions about sparsely-attended entities in the passage, to help improve neural IR models by learning to distribute attentions over the passage.

MuMuQA: Multi-Media Multi-Hop News Question Answering via Cross-Media Grounding
Under review at AAAI 2022

- Proposed a benchmark that incorporates **cross-media knowledge extraction and grounding** for multimedia question answering over news articles.

- Introduced a novel data generation framework that uses **multi-media information extraction** for generating questions that are grounded on objects in images and need to be answered using news body text.

Towards Robust Neural Retrieval Models with Synthetic Pre-Training
Under review at ECIR 2022

- Improved the **zero-shot performance** of state-of-the-art neural IR models on both **in-domain and out-of-domain** datasets by pre-training with **synthetic questions** generated automatically from raw text passages.

Research Publications (* denotes first author)

Synthetic Target Domain Supervision for Open Retrieval QA* PDF — ***SIGIR 2021 (poster)***

- Leveraged an **automatic text-to-text generation** idea to improve the performance of state-of-the-art open-domain **end-to-end** question answering systems in a specialized domain, such as COVID-19.

InfoSurgeon: Information Consistency Checking for Fake News Detection PDF — ***ACL 2021***

- Proposed a cross-media **fake news detection** system that identifies misinformation at the **knowledge element level**, with improvements in detection accuracy and better model interpretability.

Leveraging AMR for Knowledge Base Question Answering PDF — ***Findings of ACL 2021***

- Proposed a **neuro-symbolic question answering system** that leverages AMR for **question understanding** and uses a pipeline-based approach involving a semantic parser, entity linkers and a neuro-symbolic reasoner.

Multi-Stage Pre-training for Low-Resource Domain Adaptation* PDF — ***EMNLP 2020***

- Proposed **synthetic pre-training** objectives by using structure in unlabeled text, that can transfer to downstream tasks with considerable gains in the **IT Domain.**

Answer Span Correction in Machine Reading Comprehension* PDF — ***Findings of EMNLP 2020***

- Proposed an approach for **correcting partial match answers** (EM=0, 0<F1<1) into exact match (EM=1, F1=1) and obtained upto **1.3%** improvement in both monolingual and multilingual evaluation.

Pushing the Limits of AMR Parsing with Self-Learning* PDF — ***Findings of EMNLP 2020***

- Proposed **self-learning approaches** via generation of synthetic text and synthetic AMR as well as refinement of actions from the oracle, achieving **state-of-the-art** performance on benchmark AMR 1.0 and AMR 2.0 datasets.

Multi-Level Memory for Task Oriented Dialogs* PDF — ***NAACL 2019 (poster)***

- Designed a novel multi-level memory architecture that retains **natural hierarchy** of the knowledge base without breaking it down into **subject-relation-object** triples, with 15-25% improvement in entity F1.

A Formal Language Approach for Generating Graphs* PDF — ***SDM 2019 (oral, poster)***

- Proposed a graph generative model based on **probabilistic edge replacement** grammars and designed an algorithm to build graph grammars by capturing the statistically significant **sub-graph patterns**.

FigureNet: A Deep Learning model for Question-Answering on Scientific Plots* PDF — ***IJCNN 2019 (oral)***

- Designed a modular network for **visual reasoning** on scientific plots, achieving state-of-the-art accuracy on **FigureQA** dataset (Maluuba-Microsoft), bettering **Relation Networks** (Google DeepMind) by **6.96%**.

Patents

Improving Model Performance through Text-to-Text Transformation via Distant Supervision
Filed with IBM Research

Method for Answer Span Correction
Filed with IBM Research

Scholastic Achievements

- Awarded the Siebel Scholarship for the class of 2022. (Link)
- All India Rank - 127 in JEE Advanced 2014, taken by more than 1.3 million students.
- Shortlisted among the top 32 in India from engineering stream for Aditya Birla Scholarships 2014.
- All India Rank - 44 in KVPY 2012, taken by close to 200,000 students.