YANDA CHEN

420 West 116th Street, New York, New York, USA

Tel: 917-602-8678 Email: yanda.chen@cs.columbia.edu yandachen.github.io

EDUCATION

Columbia University **New York, NY** **Sep 2021 – Present**
Ph.D. in Computer Science GPA: 4.0 / 4.0

Columbia University **New York, NY** **Aug 2017 – May 2021**
B.S. in Computer Science GPA: 4.0 / 4.0

PUBLICATIONS

[1] **Yanda Chen**, Chen Zhao, Zhou Yu, Kathleen McKeown, He He. On the Relation between Sensitivity and Accuracy in In-context Learning. 2022. *ArXiv*.

[2] Yukun Huang, **Yanda Chen**, Zhou Yu, Kathleen McKeown. In-context Learning Distillation: Transferring Few-shot Learning Ability of Pre-trained Language Models. 2022. *ArXiv*.

[3] **Yanda Chen**, Ruiqi Zhong, Sheng Zha, George Karypis, He He. Meta-learning via Language Model In-context Tuning. 2022. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*.

[4] **Yanda Chen**, Chris Kedzie, Suraj Nair, Petra Galuscakova, Rui Zhang, Douglas Oard, Kathleen McKeown. Cross-language Sentence Selection via Data Augmentation and Rationale Training. 2021. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*.

[5] **Yanda Chen**, Md Arafat Sultan, Vittorio Castelli. Improved Synthetic Training for Reading Comprehension. 2020. *ArXiv.*

[6] Ruiqi Zhong, **Yanda Chen**, Desmond Patton, Charlotte Selous, Kathleen McKeown. Detecting and Reducing Bias in a High Stakes Domain. 2019. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing.*

HONORS

- Avanessians Doctoral Fellowships for Engineering Thought Leaders and Innovators in Data Science. 2023.
- Mudd Doctoral Fellowship, Columbia SEAS. 2021.
- Honorable Mention, CRA Undergraduate Research Awards. 2021.
- Theodore R. Bashkow Research Award, Columbia Computer Science Dept. 2021.

INTERNSHIP EXPERIENCE

Amazon Web Services (AWS) AI **Jun 2021 – Sep 2021**
Research Intern at NLP group

- Proposed a novel few-shot meta-learning method called in-context tuning, where training examples are used as prefix in-context demonstrations during task adaptation.
- Showed that in-context tuning out-performs MAML in terms of both accuracy and optimization stability.
- Demonstrated that in-context tuning can eliminate well-known artifacts of few-shot language model prompting such as over-sensitivity to example ordering, example selection and instruction wording.

IBM Research **Jun 2020 – Sep 2020**

Research Intern at NLP group

- Explored novel synthetic training methods under the setting of machine reading comprehension.
- Proposed targeted synthetic training where a carefully selected subset of synthetic training examples improves model performance.
- Invented synthetic knowledge distillation and showed that distillation with synthetic training examples can close the performance gap between small student models and large teacher models.

Microsoft **Jun 2018 – Aug 2018**

Premier Field Engineer at AI group

- Designed and built prediction models for a medicine manufacture company to predict future dates of machinery breakdown based on time sequences of operational machinery data.
- Applied machine learning and deep learning models including GBDT, CNN and LSTM for data analysis.
- Mitigated the label imbalance challenge with various under-sampling and over-sampling techniques
- Achieved a macro-F1 score of 0.88 on a held-out test set.

RESEARCH EXPERIENCE

On the Relation between Sensitivity and Accuracy in In-context Learning **Mar 2022 – Sep 2022**

- Found that in-context learning sensitivity is heavily obscured by label bias
- Discovered that in-context learning sensitivity is strongly and negatively correlated with accuracy, with sensitive predictions being significantly less confident
- Proposed a few-shot selective prediction method SenSel that abstains from sensitive predictions, and showed that SenSel outperforms two strong confidence-based selective prediction baselines

In-context Learning Distillation: Transferring Few-shot Learning Ability of Pre-trained Language Models **Jun 2022 – Dec 2022**

- Proposed in-context learning distillation, which transfers in-context learning (ICL) ability from large language models to small language models by augmenting in-context tuning with teacher-student distillation
- Improved the ICL performance of small language models on CrossFit with in-context learning distillation
- Further improved the transfer by combining in-context learning distillation and language modeling distillation

Cross-language Sentence Selection via Data Augmentation and Rationale Training **Jun 2019 – Feb 2020**

- Proposed a novel approach to cross-language query-focused sentence selection in low-resource settings
- Overcame the low-resource challenge with a data augmentation algorithm by leveraging noisy parallel corpus
- Improved model performance by using alignment information from a phrase-based statistical machine translation model as additional supervision
- Achieved state-of-the-art results on both text and speech across three language pairs (English-Somali, English-Swahili and English-Tagalog)

Detecting and Reducing Bias in a High Stakes Domain **Jan 2019 – Jun 2019**

- Proposed a framework to systematically detect and reduce language bias in deep learning models
- Applied the proposed framework to a state-of-the-art social media analysis model that automatically detects aggression and grief emotions in online tweets

- Detected and verified inherent language bias of the model using leave-one-out analysis and adversarial attacks
- Drastically reduced the language bias of the model by guiding model attention with human rationale

Better Lower Bound of Success Probability of Karger-Stein Algorithm **Jan 2020 – Jun 2020**

Advisor: *Clifford Stein*

- Investigated whether the success probability of Karger-Stein algorithm has a tighter lower bound than the established $O(n^{-k})$ bound in the literature if the number of minimum k-cuts in a graph with n vertices is small
- Derived an impossibility result that the success probability of Karger-Stein algorithm can be arbitrarily close to the lower bound $O(n^{-k})$ even if there is only one minimum k-cut in a graph with n vertices
- Proved an exponentially better lower bound by using sunflower lemma with an inductive argument
- Showed that the new lower bound is tight for some graphs with few small k-cuts such as n-cliques

COURSES

Natural Language Processing, Deep Learning, Computer Vision, Machine Learning, Advanced Machine Learning, Algorithms for Massive Data, Randomized Algorithms, Analysis of Algorithms, Computational Learning Theory, Graphical Models, Social Network Analysis, Real Analysis, Abstract Algebra

TEACHING EXPERIENCE

Natural Language Processing (CS 4705)

- Spring 2022: Taught by Prof. Zhou Yu
- Spring 2021: Taught by Prof. Kathleen McKeown

Analysis of Algorithms (CSOR 4231)

- Spring 2021: Taught by Prof. Clifford Stein
- Spring 2020: Taught by Prof. Eleni Drinea

SKILLS

- Programming Languages: Python, Java, MySQL, C, C++
- Tools: PyTorch, HuggingFace, TensorFlow, Keras, Microsoft CNTK, OpenCV, Android Studio