WILLIAM CHEN

[redacted] ◇ wanchichen.github.io

EDUCATION

University of Central Florida, Orlando, Florida *2018-2021*
B.S. with Honors in Computer Science, Magna Cum Laude
B.A. with Honors in History, Cum Laude
Burnett Honors College
GPA: 3.9/4.0

University of Central Florida, Orlando, Florida *2017-2018*
Dual Enrollment
GPA: 4.0/4.0

University of Florida, Gainesville, Florida *2017-2018*
Dual Enrollment
GPA: 3.69/4.0

RESEARCH EXPERIENCE

My research interests are in Natural Language Processing, primarily in the field of Machine Translation. I am specifically interested in Machine Translation for Low-Resource Languages. I want to develop computational approaches that work for all languages, especially those without massive amounts of speakers or resources, so that machine translation services are inclusive of speakers of any language and can function for a diverse user base.

University of Central Florida, Computational Biology Lab June 2020 - Present
Research Assistant

Worked with Dr. Wei Zhang and Jiao Sun to develop multi-omics models for cancer sub-type prediction. We developed a method to integrate multiple hierarchies of biological information into the model architecture using graph networks and model ensembling, allowing for better identification of biological regions of interest. We also experimented with deep-learning techniques such as transfer-learning and cross-modal attention to take advantage of multi-modal information and better utilize data for low-resource sub-types.

University of Central Florida, Evolutionary Computation Lab Jan. 2020 - Oct. 2021
Research Assistant

Worked with Dr. Annie Wu to explore the applications of cellular automata to enhance file compression algorithms. I developed a environment to visualize the evolution of cellular automata under a variety of conditions. We utilized this information to devise an entropy-based generative model model that could create large amounts of synthetic data for the evaluation of file compression algorithms, even with a small training corpus.

University of Central Florida, Security and Analytics Lab April 2021 - July 2021
Research Assistant

Worked with Dr. David Mohaisen on applied NLP in cybersecurity. We developed a multi-task learning framework to better enable large-scale language models to extract technical data and summarize public software vulnerability reports into a structured format with only 1000 samples of domain-specific text.

WORK EXPERIENCE

Texas Instruments July 2021 - Present
Software Engineer *Dallas, TX*

Full-stack developer on the E-commerce Inventory Management Team, working in Java Spring, React.js, and SQL. Primary product owner and developer of the e-commerce portal, a web application to view and control the one billion dollars of order flow and inventory in the entire e-commerce environment.

uBump.co Aug. 2020 - May 2021
Chief Information Officer *Orlando, FL*

Lead front-end and UI development for the flagship website of an NFC-oriented social media sharing startup. Our startup was acquired by Bolstered Equity Group in May 2021.

Valorantify June 2020 - Aug. 2020
Software Engineer

Front-end developer for one of the first VALORANT e-sports sites, which became one of the top statistics and news sources for the professional scene. Acquired by thespike.gg.

Texas Instruments June 2020 - Aug. 2020
Software Engineering Intern *Dallas, TX*

Worked with a global team of developers and stakeholders to build a React dashboard and Java Spring API server to track and manage orders throughout the supply chain pipeline, improving worker allocation throughout product distribution centers worldwide.

FUNDING, AWARDS AND HONORS

FLORES 101 Compute Grant (2021)	$750 in computing credits from Facebook Research
NSF Research Grant (2020)	Funded work at UCF Computational Biology Lab
Benaquisto Scholarship (2018)	Full-ride merit scholarship funded by Florida
National Merit Finalist (2018)	Awarded to top 1% of PSAT scorers

PUBLICATIONS

William Chen and Brett Fazio. **Morphologically-guided Segmentation for Translation of Low-Resource Agglutinative Languages**. *Proc. of the Workshop on Technologies for Machine Translation of Low Resource Languages (LoResMT), 2021.* (**Best Paper Award Honorable Mention**.)

William Chen and Brett Fazio. **The UCF Systems for the LoResMT 2021 Machine Translation Shared Task**. *Proc. of the Workshop on Technologies for Machine Translation of Low Resource Languages (LoResMT), 2021.*

Khandakar Tanvir Ahmed, Jiao Sun, William Chen, Irene Martinez, Sze Cheng, Wencai Zhang, Jeongsik Yong, and Wei Zhang. **In Silico Model for miRNA-mediated Regulatory Network in Cancer**. *Briefings in Bioinformatics, Volume 22, Issue 6, 2021.*

UNPUBLISHED MANUSCRIPTS

Sukirna Roy, William Chen, and Annie S. Wu. **Central Florida Corpus: A Generative Synthetic File Corpus**. *In Review.*

William Chen, Brett Fazio, and David Mohaisen. **Multi-task Learning for Normalization and Summarization of Software Vulnerability Reports**. *In Preparation.*

William Chen, Kensal Ramos, Kalyan Naidu Mullaguri, and Annie S. Wu. **Genetic Algorithms for Extractive Summarization**. *arXiv:2105.02365, 2021.*

RELEVANT PROJECTS

Multilingual Machine Translation Model
Developed a multilingual machine translation model for the WMT 2021 FLORES 101 Multilingual Machine Translation Shared Task (Southeast Asian Track). Experimentation focused primarily on exploiting the similar morphology of the languages, using techniques such as model ensembling and transfer learning. Obtained initial results, but was not able to submit a final model due to computing constraints.

Taiwanese Sign Language Translation Model
Designed a pipeline for translation of Taiwanese Sign Language video clips to Traditional Chinese text, adapting the TimeSformer variant of the Vision Transformer for a generative video to text model. Project was initially for the LoResMT 2021 Shared Task, although was not submitted due to unsatisfactory results.

Basketball Tracker for Video
Created a model for tracking basketballs in video for an end-of-year project of a graduate-level Computer Vision course. We combined a pre-trained YOLO model for bounding box detection with Discriminative Correlation Filters for live tracking, even without a GPU.

SKILLS

Languages	English (native), Mandarin Chinese (native), French (fluent)
Programming Languages	Python, Java, Javascript, Typescript, C, C#, Rust
Frameworks	PyTorch, Tensorflow, React, Express, Actix-Web

SERVICE

ACL-IJCNLP 2021 Micro-blogging Volunteer

REFERENCES

Dr. Annie S. Wu, Associate Professor
Department of Computer Science
University of Central Florida, Orlando FL


Dr. David Mohaisen, Associate Professor
Department of Computer Science
University of Central Florida, Orlando FL


Dr. Wei Zhang, Assistant Professor
Department of Computer Science
University of Central Florida, Orlando FL