

# PINQIAO WANG

## CURRICULUM VITAE

### Education

#### University of North Carolina at Chapel Hill

2019 – 2023

*B.S. in Statistics and Analytics, Double Major in Economics, Major GPA: 3.88*

*Chapel Hill, NC*

- **Thesis:** Using Unsupervised Learning Methods for Sentiment Analysis

#### Columbia University

2023 – 2024

*M.A. in Statistics*

*New York, NY*

### Relevant Coursework

- **Math:** Linear Algebra, Discrete Math, Multivariate Calculus, Differential Equations, Probability Theory, PhD-level Real Analysis, Mathematical Analysis/Measure Theory, etc.
- **Statistics:** Statistical Inference, Optimization, Machine Learning, Stochastic Process, Statistical Analysis Methods, PhD-level Theoretical Statistics, etc.
- **CS/DS:** Data Structures and Algorithms, Data Science, Deep Learning/Neural Networks, Data Mining, NLP, LLM, etc.

### Research Interests

- AI4Science (Material/CarbonTech/Medical/Finance)
- Efficient/Trustworthy ML/AI
- LLM Agents
- Computer Vision
- Natural Language Processing

### Research Experience

#### Research Assistant | Supervisor: Dr. Arain Maleki | Columbia Stats Dep.

Feb 2024 – Aug 2024

*Columbia University*

*New York, NY*

- Developed custom neural network models tailored for high-dimensional imaging data, utilizing PyTorch and TensorFlow to enhance computational efficiency and achieve a 7 percent increase in accuracy on benchmark datasets.
- Created and fine-tuned 2 LSTM models for image segmentation and classification, demonstrating significant improvements in precision and recall metrics, essential for medical and remote sensing applications.
- Integrated KNN imputation and LLM to successfully address/synthesize noisy/incomplete data challenges.

#### Research Assistant | Supervisor: John Cornwell/Dr. Nicolò Daina | CDI

Feb 2024 – May 2024

*Columbia University, Center on Global Energy Policy*

*New York, NY*

- Assisted in designing machine learning algorithms to optimize carbon capture processes, utilizing PyTorch and TensorFlow to model complex chemical interactions and predict optimal operating conditions, achieving a 30 percent increase in efficiency.
- Built and fine-tuned LLMs with LLaMA2 to automate literature reviews and generate research hypotheses in the field of carbon tech, significantly reducing the time required for data analysis and hypothesis generation.
- Preprocessed and analyzed extensive datasets related to carbon emissions and climate impact, employing techniques such as normalization, feature extraction, and dimensionality reduction to ensure data quality and model accuracy.

### Professional Experience

#### Agam Capital Management, LLC

Jun 2024 – August 2024

*AI Research Intern*

*New York, NY*

- Focus on customizing and fine-tuning Large Language Models. Design RAG + MoE machine inference pipeline to achieve 85 percent accuracy in automated code translating and Q/A generation.
- Prototype the latest AI/ML research and localize the method tailored to one's own business needs.

#### AI4Finance Foundation

Oct 2023 – present

*NLP/LLM Intern part-time*

*New York, NY*

- FinGPT Research (Focused on FinRobot): Financial LLM framework with 10000+ stars on [GitHub](#).
- Applied RAG to engineer scalable pipelines to reduce hallucinations and conduct inference. Model Training for stock forecasting in the Bitcoin market, applied SWT and LoRA in Python to fine-tune FinGPT with A100 GPU with Llama2 and Baichuan2, reached 70 percent accuracy; Optimized NLP algorithms to enhance FinRobot output for auto financial report generation.

Projects

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| <b>CDSS Hackathon Winner Project</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | <b>Oct 2023</b>     |
| <i>Team Member</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | <i>New York, NY</i> |
| <ul style="list-style-type: none"><li>• Collaborated with a team of 4 to create a framework targeting energy inefficiency in NYC buildings, challenging existing energy efficiency rating systems by identifying anomalies using Python. Here is the <a href="#">code</a>.</li><li>• Co-designed a Bayes hierarchical model using Bayesian Lasso for variable selection and detected "Local Outliers" Achieved 1st place out of 107 participants in the competition funded by HRT and Google Cloud.</li></ul> |                     |

Honors and Awards

|                                           |                                                    |
|-------------------------------------------|----------------------------------------------------|
| <b>Dean’s List</b>                        | <b>2021 – 2023</b>                                 |
| <i>Award for Academic Excellence</i>      | <i>University of North Carolina at Chapel Hill</i> |
| <b>Student Representative</b>             | <b>2023</b>                                        |
| <i>Departmental Award</i>                 | <i>Columbia University</i>                         |
| <b>MCM/ICM 2024 M award</b>               | <b>2024</b>                                        |
| <i>Columbia University</i>                | <i>MCM</i>                                         |
| <b>Fu Foundation Data Science Scholar</b> | <b>2024</b>                                        |
| <i>Data Science Insitute</i>              | <i>Columbia University</i>                         |

Technical Skills

**Languages:** Python, R, SQL, C, C++, C-sharp, SAS, Power BI, MATLAB, Microsoft Office, Google Cloud, AWS  
**Technologies/Frameworks:** Linux, GitHub, WordPress