Prince Osei Aboagye

Curriculum Vitae

Visa Research
385 Sherman Ave
Palo Alto, California, 94306
915-867-2874
prince@cs.utah.edu
poaboagye.github.io

Research Interests

My research endeavors to propel Artificial Intelligence (AI) advancement, explicitly targeting the enhancement of cross-lingual adaptability in language technologies, especially for under-resourced languages. My investigative scope encompasses Natural Language Processing, Speech Representation Learning, and Knowledge Representation Learning, extending into monolingual, cross/multilingual, and multi-modal situations. Leveraging my prior work on disentangling conceptual associations in word vector embeddings to improve interpretability, my current focus is on advancing this concept for Multi-Class Classification and Out-of-Distribution Detection tasks. By fine-tuning the embedding function to allocate unique coordinates to each class, thus fostering orthogonal representations, my research aims to bolster the interpretability, structure, and accountability of automated decision-making processes. The implications of this research are directly tied to large language models, aiming not only to enable and facilitate their practical usage but also to enhance their reliability and adaptability to diverse needs and contexts, thus ensuring a more responsible and ethical application of AI systems.

Experience

2023 - **Visa Research**, *Staff Research Scientist*

Education

2018 – 2023 **PhD**, *Computer Science*, University of Utah

Advisor **Dr. Jeff Phillips**

2016 – 2018 **MS**, *Applied Mathematics*, University of Texas at El Paso

Advisor **Dr. Michael Pokojovy**

2011 – 2015 **BA**, *Mathematics and Economics*, University of Ghana

Advisor **Mrs. Lilian Frempomaa Kyei**

Technical Skills

- Proficient in machine learning and deep learning for multiple applications
- Python, Java, C++, MatLab, Tensorflow, Pytorch

Patent

- Zheng, Yan, Michael Yeh, Wang Junpeng, Wei Zhang, Liang Wang, Hao Yang, and **Prince Osei Aboagye**. "Method, System, and Computer Program Product for Normalizing Embeddings for Cross-Embedding Alignment." U.S. Patent Application 18/006,649, filed July 6, 2023.
  `https://patents.google.com/patent/US20230214177A1/en`

Publications

Under Review 2023: Facilitating Image-Text Comprehension with Visual Analytics. Yiran Li, Junpeng Wang, **Prince Osei Aboagye**, Michael Yeh, Yan Zheng, Liang Wang, Wei Zhang.

Under Review 2023: PDT: Pretrained Dual Transformers for Time-aware Bipartite Graphs. Xin Dai, Yujie Fan, Zhongfang Zhuang, Shubham Jain, Chin-Chia Michael Yeh, Junpeng Wang, Liang Wang, Yan Zheng, **Prince Osei Aboagye**, Wei Zhang. **Link to paper**: `https://arxiv.org/abs/2306.01913`

Under Review 2023: One-Hot Encoding Strikes Back: Fully Orthogonal Coordinate-Aligned Class Representations. **Prince Osei Aboagye**, Hasan Pourmahmoodaghababa, Yan Zheng, Chin-Chia Michael Yeh, Junpeng Wang, Huiyuan Chen, Xin Dai, Liang Wang, Wei Zhang, Jeff Phillips.

ICLR 2023: Interpretable Debiasing of Vectorized Language Representations with Iterative Orthogonalization. **Prince Osei Aboagye**, Yan Zheng, Jack Shunn, Chin-Chia Michael Yeh, Junpeng Wang, Zhongfang Zhuang, Huiyuan Chen, Liang Wang, Wei Zhang, Jeff Phillips **Link to paper**: `https://openreview.net/forum?id=TkQ1sxd9P4`

ICLR 2022: Normalization of Language Embeddings for Cross-Lingual Alignment. **Prince Osei Aboagye**, Jeff Phillips, Yan Zheng, Junpeng Wang, Chin-Chia Michael Yeh, Wei Zhang, Liang Wang, Hao Yang. **Link to paper**: `https://openreview.net/forum?id=Nh7CtbyoqV5`

AMTA 2022: Quantized Wasserstein Procrustes Alignment of Word Embedding Spaces. **Prince Osei Aboagye**,Yan Zheng, Michael Yeh, Junpeng Wang, Zhongfang Zhuang, Huiyuan Chen, Liang Wang, Wei Zhang, and Jeff Phillips. **Link to paper**: `https://aclanthology.org/2022.amta-research.15/`

Research Experience

Fall 2022- **Current Graduate Research**, *School of Computing and Visa Research*, University of Utah

Advisor: Dr. Jeff Phillips

Description: Models that jointly predict task labels and generate free-text explanations for their predictions, also known as self-rationalization models are greatly interested in modern Explainable AI. This leads to a more intuitive interaction with NLP systems. Given the free-text explanations from Self-rationalization models (say a GPT-3 model), our goal is to express or assign an uncertainty score to the generated free-text explanations to understand the level of confidence that the Self-rationalization model places on its own generated free-text explanations.

Summer 2022 **Research Scientist Intern**, Visa Research

Advisor: Wei Zhang

Title: Concept Denoising of Convolutional Neural Networks with Spectral Normalization and Iterative Orthogonalization

Spring 2022 **Graduate Research**, *School of Computing and Visa Research*, University of Utah

Advisor: Dr. Jeff Phillips

Description: While existing methods mitigate language representation biases, they only work for just two identified concepts (such as gender and occupation or race and religion). Hence, they can't be used to debias multiple concepts simultaneously. Also, the approaches employed by these existing methods are too aggressive: they not only remove bias but also erase valuable information from word embeddings, which turn to impact downstream task performance negatively. Following this, we are developing a method that simultaneously decouples or othogonalizes multiple concepts instead of removing concepts wholesale so that semantic information is retained in the embeddings and bias is also effectively mitigated compared to other existing methods.

Summer 2021 **Research Scientist Intern**, Visa Research

Advisor: Wei Zhang

Title: Quantized Wasserstein Procrustes Alignment of Word Embedding Spaces

Description: We proposed an unsupervised cross-lingual word embedding (CLWE) model that poses the alignment task as a Wasserstein-Procrustes problem to jointly estimate a permutation matrix and an orthogonal matrix . I relied on a quantization step to find the closest measure supported on k points in the 2-Wasserstein distance through Optimal Transport (OT) methods. This substantially improves the approximation quality of empirical OT solvers, given fixed computational cost. Our proposed unsupervised CLWE model obtains state-of-the-art results on the Bilingual lexicon Induction (BLI) task.

Fall 2020-Spring 2021 **Graduate Research**, *School of Computing and Visa Research*, University of Utah

Advisor: Dr. Jeff Phillips

Description: In this work, we proposed a new and general approach to normalize word embeddings that improves the level of isomorphism between monolingual embedding spaces. The key is Spectral Normalization which regularizes the spectral properties of monolingual embeddings. Moreover, we show layering Spectral Normalization within an iterative sequence with also centering and vector length normalization improves results further. We demonstrate this improvement on the standard bilingual lexicon induction (BLI) and Cross-lingual downstream tasks (CLDTs).

Fall 2017-Spring 2018 **Master Thesis**, *Department of Mathematical Sciences*, University of Texas at El Paso

Advisor: Dr. Michael Pokojovy

Title: On Numerical Stochastic Optimal Control Via Bellman's Dynamic Programming Principle

Goal: In this work, we presented an application of Stochastic Control Theory to the Merton's portfolio optimization problem. Then, the dynamic programming methodology was applied to reduce the whole problem to solving the well-known HJB (Hamilton-Jacobi-Bellman) equation that arises from the Merton's portfolio optimization problem subject to the power utility function. Finally, a numerical method was proposed to solve the HJB equation and the optimal strategy. The numerical solutions are compared with the explicit solutions for optimal consumption and investment control policies.

2014 – 2015 **Undergraduate Thesis**, *University of Ghana*, Mathematics Department

Advisor: Ms. Lilian Frempomaa Kyei

Description: I got acquainted with Game Theory, the definition of games in the normal form as well as the matrix representation of two-player finite games. I looked at the famous prisoner's dilemma game and how the conflict between the social incentive to cooperate and the private incentive to defect can be resolved

Previous Projects

CS 6350: Fall 2018 **Machine Learning**, *Kaggle: Classifying Movie Reviews*, University of Utah

Description: Each example in this classification task is a movie review. The goal was to predict whether the review was positive or negative . The data used for this task is based on the Large Movie Review Dataset v1.0. Each review in this task is characterized by the histogram of the words it contains. I implemented the following machine learning models for this classification task: Simple Perceptron, Average Perceptron, Support Vector Machines, Logistic Regression, Bagging on Support Vector Machines, and AdaBoost.

CS6955: Fall 2018 **Deep Learning**, *Classifying Movie Reviews Using Recurrent and Recursive Neural Network*, University of Utah

Description: Sentiment classification over the past decades has been solved using linear classification methods, such as Support Vector Machines (SVM), Random Forest, and logistic regression. Even though the above models may result in reasonable accuracies, they all suffer from losing the order of words appearing in a sentence. They, therefore, cannot capture delicate semantics from the input reviews. Therefore, we resorted to Recursive and Recurrent neural networks to allow us to account for the order of words in a sentence.

CS 6170: Spring 2019 **Computational Topology**, *Sentiment Classification with Topological Signatures*, University of Utah

Description: A common approach in Topological Data Analysis (TDA) is to capture the shape or the underlying structure of shapes in data. Of course, it is not easy to define meaningful shapes in textual documents. However, when text is interpreted as describing a progression of events (as in movies), topological features, namely, homological persistence, can significantly improve classification accuracy when added to a text vector representation.

CS 6956: Spring 2019 **Deep Learning for Natural Language Processing**, *Music Genre Classification By Lyrics Using Deep Neural Network* , University of Utah

Description: Music genre classification based on lyrics alone is a very important and heavily researched task in Music Information Retrieval. The ever-growing amount of online music databases calls for intelligent tools to help people browse and search these music databases and categorize and organize the songs they listen to. For this project, we explored several machine learning algorithms to identify and classify the genre of a song given its lyrics.