# **GEORGE HALAL** | Personal Site: <a href="https://georgehalal.github.io">https://georgehalal.github.io</a> | Email: <a href="halalgeorge@gmail.com">halalgeorge@gmail.com</a> | Phone: +1 (650) 422-9033

**EDUCATION** 

Stanford UniversityPh.D. PhysicsGPA: 4.00/4.002019–2024Lehigh UniversityB.S. Physics & Minor in Applied MathematicsGPA: 3.97/4.002015–2019

**SKILLS** 

Python Packages | PyTorch • PyTorch Lightning • vllm • WandB • asyncio • Pandas • Hugging Face (transformers, tokenizers,

datasets, evaluate, accelerate, peft, trl) • NumPy • Statsmodels • SciPy • Seaborn • Xgboost • Shap •

Matplotlib • Requests • RE • Scikit-learn

Other Languages/Tools | C++ • SQL • HTML • MATLAB • SLURM

Al Topics | LLMs/Natural Language Processing • Computer Vision • Sequence Modeling • MLOps
Other Topics | Causal Inference • Bayesian Inference • A/B Testing • Time Series Analysis/Signal Processing

#### **EXPERIENCE**

### Member of Technical Staff | Contextual AI, Mountain View, CA

| 2024

#### **Graph-based Retrieval Augmented Generation for Enterprise LLMs**

Developing state-of-the-art graph-based retrieval augmented generation (GraphRAG) solutions for large language models (LLMs)
in enterprise settings, from research to production.

## Data Scientist Intern | Alife Health, San Francisco, CA

| 2023

#### Causal Inference and Machine Learning for IVF Intracycle Dose Adjustments

- Developed techniques for analyzing the impact of dose adjustment patterns throughout IVF cycles on pregnancy outcomes.
- Employed statistical tests to alert clinics when a doctor's performance deviates from average on key performance indicators (KPIs).

# Graduate Student Researcher | Stanford University, Stanford, CA

2019-2024

#### Transformer-Based Super-Resolution for Dust Polarization Images | GitHub Link

- Built a multi-image encoder, a transformer-based fusion module, and a decoder to increase the image resolutions by 4x.
- NeurIPS paper in prep.

#### Causal Inference for Modeling the Effects of the Nearby Dust Geometry on Magnetic Fields | Paper Link

#### Spherical Harmonic Convolutional Hough Transform | GitHub Link | Paper Link | Invited Talk Link

- Developed a computer vision algorithm to model the structure of interstellar gas.
- Achieved 3000x runtime speedup and 5x memory reduction over the previous state-of-the-art.

# Modeling the Foreground Obscuring Radiation from the Early Universe | Paper Link | Award Link | Invited Talks: Harvard, Spain, S4

Used computer vision, hypothesis testing, and Bayesian inference for quantifying this foreground signal, setting new limits.

#### Deep Learning for Stochastic Generation of Observed Galaxy Properties | GitHub Link

• Developed a conditional Wasserstein generative adversarial neural network with gradient penalty (cWGAN-GP) to generate observed galaxy properties in wide-field surveys.

# Deep Learning for Modeling the Transfer Function of Galaxy Detection | GitHub Link

 Developed a probabilistic model for predicting the transfer function of galaxy detection in wide-field surveys, achieving an ROC-AUC score of 0.95.

# Deep Learning for Searching for 2- $\nu$ Double- $\beta$ Decay of <sup>136</sup>Xe | Poster Link

Developed a Long Short-Term Memory (LSTM) based model to search for this decay to the excited state of <sup>136</sup>Ba in EXO-200 data, achieving an ROC-AUC score of 0.98.

## **Undergraduate Student Researcher** | Yale University and Lehigh University

| 2018-2019

#### Deep Learning for Heavy-Flavor Jet Classification at RHIC | Report Link | Talk Link

• Developed a Python-based Long Short-Term Memory (LSTM) model to classify bottom, charm, and light jets, attaining misclassification rates of 2.1%, 10.9%, and 4×10<sup>-3</sup>%, respectively, leveraging C++ for efficient data preprocessing.

#### Undergraduate Student Researcher | The Ohio State University and Lehigh University

2016-2018

## **Deep Learning for Collision Geometry Determination**

• Developed a model to identify the collision geometry of nuclei based on the activation pattern of STAR-EPD detector tiles in Python, leveraging C++ for efficient data preprocessing.