Graduate Plan:

- Intellectual Merit: Harnessing wisdom of an artificial crowd tom improve LLM Reasoning Capability
  - Value gained in Fermi Estimations
- Currently researchers are trying to improve reasoning using the following techniques
  - Self-consistency
  - Conformal Prediction
  - Ensemble methods
    - Transition to Project
- Hypothesis & Goal
  - Law of Large Numbers should improve estimations when calculating answers to complex tasks
  - To approach problem, I will utilize HPC to test different implementations of LLN applied to LLM Reasoning
  - Hypothesize that LLN applied to estimating components of output should construct the most accurate response
- Include a figure/diagram that represents the construction of certain tests/experimental setups
- Approach
  - Aim 1 — Benchmark: Build and release a curated, unit-normalized Fermi benchmark with authoritative sources and acceptable ranges.
  - Aim 2 — LLN: Implement Law of Large Numbers in LLM reasoning tasks.
  - Aim 3 — Methods:
    - (3a) Do LLM ensembles exhibit a predictable reduction in Fermi estimation error with more diverse samples? What is the convergence

rate of log-error as a function of sample size and diversity (temperature, prompts, model families)?.

- (3b) Can we elicit and calibrate distributional forecasts (e.g., 50% and 90% intervals) with correct coverage on Fermi tasks?

- (3c) Can structured factorization with unit-safe computation and retrieved base rates reduce systematic log-error and overconfidence?

- Aim 4 — Evaluation
- Broader Impacts
  - Expand the idea of harnessing LLN to improve Reasoning
  - Better LLM estimations can support LLM ensembles/checking
  - Connect to past work done in internship/research experience/project

Personal Statement:

- Commitment to service
    - SafeWalk App
    - Habitat for Humanity
    - Kaiser Permanente
    - Animal Shelter
- Research Lab Reading Groups
    - Fluids Lab
        - Lorenz-96 Two timescale system analysis on weather simulations
    - Robotics Lab
        - Imitation Learning Algorithm analysis
        - RL
- Research Work
    - Palo Alto Networks
        - ML Unstructured vs. Structured UPD
            - Production MaaS Push
            - PII Security
        - LoRA Model Hardening Transferability
    - L3 Lab:
        - Background/Motivation
            - AI + Deep Learning Class
            - Reading papers
        - LoRA Model Hardening Transferability Experimentation
            - UNC Undergrad Research Conference
        - LLN Galton Project
            - Experiments being run

- - - Papers presented to Lab
  - USG
    - Data Engineering
    - Pitched Novel solution and improved pipeline
  - CS TA
    - Education Research
  - Chancellor Science Scholar
    - 2x Distinguished Scholar

Personal, Relevant Background and Future Goals Statement