In The Name of God

Final Project – Probabilities and Statistics

Fall 03 – Dr. Jamshidi

Project Overview

This final project is designed to apply the concepts learned in this course to real-world, practical problems. Students will select one abstract topic from the list below, conduct independent research, and implement a solution using the tools and methods covered in class. Students are encouraged to use their knowledge creatively and to enhance their approach by leveraging AI tools, such as ChatGPT, for problem-solving, code generation, and conceptual understanding.

The project emphasizes critical thinking, abstraction, and innovation. It is designed to challenge students to handle unknown problems, demonstrating their ability to apply theoretical knowledge to practical scenarios.

Topics

Students must choose one of the following topics:

1. Wireless Communication Channel Modeling
2. Machine Learning: Feature Selection and Prediction
3. Traffic Flow Analysis in Smart Cities
4. Sensor Fusion in IoT Systems
5. Queueing Systems in Networking
6. Energy Harvesting Systems
7. Reliability Analysis of Electrical Circuits
8. Stock Market Analysis
9. Marketing Campaign Analysis

General Guidelines

1. **Abstract Problem-Solving:** Students are expected to independently research the topic, formulate the problem, and solve it using the probabilistic tools learned in class.
2. **AI Integration:** Use AI tools like ChatGPT to assist in breaking down the problem, generating solutions, debugging code, and providing additional learning resources. Clearly document how AI tools were utilized.
3. **Deliverables:** Each project must include:
    - Problem formulation and methodology

- Mathematical modeling and calculations
- Simulation or data analysis
- Visualization (e.g., plots, graphs, tables)
- Discussion and conclusion

Project Topics and Guides

1. Wireless Communication Channel Modeling

- **Objective:** Model a wireless communication channel using probabilistic methods. Analyze signal fading and interference.
- **Steps:**
    1. Generate random variables representing signal strength, noise, and interference (e.g., Gaussian, Rayleigh distributions).
    2. Derive the probability density function (PDF) and cumulative distribution function (CDF) for the channel output.
    3. Calculate moments (mean, variance) of the output.
    4. Simulate channel fading and plot its effects on signal quality.
    5. Analyze the impact of noise on the bit-error rate (BER) using Monte Carlo simulations.

2. Machine Learning: Feature Selection and Prediction

- **Objective:** Analyze a dataset and predict an outcome using regression models. Identify the most influential features probabilistically.
- **Steps:**
    1. Select a dataset (e.g., from Kaggle or UCI Machine Learning Repository).
    2. Calculate covariance and correlation matrices to identify relationships among features.
    3. Use regression (e.g., linear regression, MMSE) to model the outcome variable.
    4. Evaluate the model's performance using probabilistic metrics (e.g., R-squared, mean squared error).
    5. Visualize the results, including feature importance.

3. Traffic Flow Analysis in Smart Cities

- **Objective:** Model and analyze traffic flow at a busy intersection using probabilistic methods.

- **Steps:**
    1. Define random variables for vehicle arrivals and signal timings (e.g., Poisson and uniform distributions).
    2. Simulate traffic flow using Monte Carlo methods.
    3. Calculate the average waiting time and queue length for vehicles.
    4. Use probabilistic inequalities (e.g., Markov or Chebyshev) to bound waiting times.
    5. Visualize traffic patterns and propose optimizations.

4. Sensor Fusion in IoT Systems

- **Objective:** Combine noisy sensor data to estimate an unknown parameter using probabilistic techniques.
- **Steps:**
    1. Generate synthetic sensor data with added Gaussian noise.
    2. Use joint probability distributions to model sensor data fusion.
    3. Apply MMSE or weighted average techniques for parameter estimation.
    4. Calculate the variance of the estimate and compare it with individual sensors.
    5. Visualize the fused data and its reliability.

5. Queueing Systems in Networking

- **Objective:** Analyze the performance of a single-server queue in a networking scenario.
- **Steps:**
    1. Define arrival and service times as random variables (e.g., exponential distributions).
    2. Calculate key metrics such as average queue length, waiting time, and server utilization.
    3. Simulate the queueing process using Monte Carlo methods.
    4. Compare simulated results with theoretical results from queueing theory.
    5. Visualize the queue behavior under varying loads.

6. Energy Harvesting Systems

- **Objective:** Model an energy harvesting system and analyze its performance under stochastic conditions.
- **Steps:**

1. Define random variables for harvested energy and consumption rates (e.g., exponential or uniform distributions).
2. Calculate the probability of energy outage.
3. Simulate the energy storage process over time.
4. Use the law of large numbers (LLN) to analyze long-term energy availability.
5. Visualize the energy dynamics.

7. Reliability Analysis of Electrical Circuits

- **Objective:** Evaluate the reliability of a simple electrical circuit using probabilistic methods.
- **Steps:**
    1. Define random variables for component lifetimes (e.g., exponential distributions).
    2. Calculate the overall reliability of series and parallel configurations.
    3. Simulate failure events and calculate mean time to failure (MTTF).
    4. Use probabilistic inequalities to bound the failure probability.
    5. Visualize reliability curves.

8. Stock Market Analysis

- **Objective:** Analyze stock price movements using probabilistic and statistical tools.
- **Steps:**
    1. Obtain historical stock data.
    2. Model price changes as random variables (e.g., Gaussian distribution).
    3. Calculate and interpret statistical metrics such as volatility and correlation.
    4. Use regression to predict future prices and evaluate its performance.
    5. Visualize price trends and probabilistic forecasts.

9. Marketing Campaign Analysis

- **Objective:** Evaluate the effectiveness of a marketing campaign using probabilistic analysis.
- **Steps:**
    1. Obtain or simulate data on campaign reach and conversions.
    2. Model conversions as a Bernoulli random variable.
    3. Calculate metrics like expected conversion rate, variance, and confidence intervals.
    4. Use regression to analyze factors influencing conversions.

5. Visualize campaign performance and propose optimizations.

Creativity and AI Usage

Students are expected to creatively solve their chosen problem and document how AI tools like ChatGPT were used to:

- Break down complex concepts and approaches.
- Generate and debug code.
- Interpret and analyze results.
- Explore additional resources for deeper understanding.

Note: The use of AI tools should not replace understanding but complement the student's approach. Solutions should reflect individual creativity and problem-solving ability.

Submission Requirements

1. **Report:** A detailed report (5-10 pages) covering all deliverables.
2. **Code:** Well-documented and functional code in Python or MATLAB.
3. **Presentation:** A short (5-minute) presentation summarizing the problem, approach, and results.
4. **AI Usage Documentation:** A section detailing how AI tools were used.

Evaluation Criteria

1. Problem Formulation: 20%
2. Mathematical Modeling: 20%
3. Simulation and Results: 30%
4. Creativity and AI Integration: 20%
5. Presentation and Clarity: 10%