1.1 Topic Brainstorming

Starting a project when it is very open and broad can be daunting - brainstorming and getting into the right mindset is a great place to start. To begin, start by brainstorming potential topics and issues that interest you when thinking about researching for your MTFC proposal. Think of topics that you are passionate about, that directly affect your neighbors or your region, have shown up in news articles, or you are just curious to know more about. You will eventually settle on ONE topic that depends on the data sources you are able to find and the research questions you seek to answer.

Think of 3 social/societal issues topics that interest you:

1. Poverty
2. World Hunger
3. Racism

Think of 3 environmental topics that interest you:

1. Global warming from Fossil Fuels
2. Plastic Pollution
3. Electronic Waste

Think of 3 health/psychology topics that interest you:

1. Cancer
2. Depression
3. Alzheimer's

Think of 3 topics of your choice that interest you that haven't fit into any other category or overflow:

1. Education
2. AI
3. Sports

1.2 Identifying Risk

This brainstorm will be to identify any risks (a potential loss). A close reading of page 11 of The Actuarial Process Guide describes “risk” in more detail: It isthe chance that an undesirable event will occur with some value (often financial) attached to the event.

In the table below, choose 3 of your brainstormed topics.

- Brainstorm any and all possible risks (chance for loss) for that particular topic.
- Brainstorm any and all parties who are at risk.
- Brainstorm possible risk mitigation strategies (insurance, behavior changes, modifying outcomes). This part of the brainstorming might entail some basic introductory research as you think about topics more deeply

Topic	Possible Risks	What parties are at risk?	Risk Mitigation Strategies
Sports	- Injuries - CTE - Sports betting	- Players - Health insurance - Families of players - Teams - Betters	- Protective equipment
Racism	- Police brutality - Inequality of opportunity in education and societal institutions - Inequality of opportunity in economic institutions	- Minorities - Colleges - Businesses - Sociopolitical institutions	- Diversity and inclusion initiatives - Educate the youth - Don’t discriminate
Cancer	- Dying - Losing hair - Losing body parts	- Patient - Health Insurance	- Chemotherapy - Early detection
Agriculture Deterioration Due To Climate Change	- Desertification of land - Insurance Policies derived by climate change Initiatives	- Farmers - People who need to eat - People who need proper insurance policies	- In-depth analysis to provide better insurance policies - Preservation of land (dikes, conservations, etc)

1.3 Topic Research Narrow-down, Sources, Research Questions

Now that you've had a chance to identify some risks and risk mitigation strategies for a few topics, you likely have a topic that is standing out to you more than others. Take ONE of the topics and delve a little deeper. The goal now is to look for good background research and possible driving research questions.

Topic: Racial Bias in AI
Directions: Find at least 3 sources with credible research, information, data, etc. related to your topic. Questions to consider for each source: ● What kind of information is in this source? ● How much data is available? Is there a sufficient time frame or history to the data? ● Does this source address frequency or severity (or both)? You'll need at least one source addressing frequency and one addressing severity to be able to adequately quantify risk. Source #1: Predictive policing algorithms are racist. They need to be dismantled ● Source link: https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/ ● Summary: The misuse of predictive policing tools and the manipulation of data by police is a big problem. These tools, including location-based and person-based algorithms, have been shown to perpetuate systemic racism due to biased data, especially arrest data. This bias can result in the disproportionate targeting of minority communities. Source #2: Health care artificial intelligence gets biased data creating unequal care ● Source link: https://news.umich.edu/health-care-artificial-intelligence-gets-biased-data-creating-unequal-care/ ● Summary: Nicholson Price, Healthcare Policy & Innovation, contemplates that Artificial

Intelligence (AI) has proposed many more harms than benefits because of the structural flaws within the software. Due to the bias data sets that have been analyzed in the past, there are a significant portion of individuals' mental health data and other data that are inefficient. Anti-minority data has been widely distributed to the point that there an immense amount of data breaches and data insecurity that are commencing with the field, harming the overall efficiency and effectiveness of these models. AI models, at least those that are used within this specific models, will result in significant amount of problems to arise as society focuses on other problems at hand; however, it is inevitable that there will be slight amount of people who are profiled incorrectly. Boiling down to the specifics of the problems at hand, it is vital that healthcare and AI are implemented together, to increase the overall efficiency. Ultimately, it boils down to the diversity that can be accrued by this algorithm as it analyzes more sets of data from the Status Quo.

Source #3: MIT Technology Review

● Source link: https://www.technologyreview.com/2019/10/17/75285/ai-fairer-than-judge-criminal-risk-assessment-algorithm/

● Summary: AI is increasingly used to judge people, from predicting risk to recommending hires. Proponents say it can help eliminate implicit bias, but this may not be true. For example, the risk-assessment AI COMPAS, used to advise judges in the courtroom by predicting the risk of rearrest, tends to give higher risk scores to black people as opposed to white, leading to a higher proportion of black people being needlessly jailed. It is difficult to balance this out without conflicting with other definitions of fairness; for example, if the threshold to release is different for different races, even if the proportion of needless jailing is the same, this seems to hold black and white people to different standards on the same scale. This is because of the inherent bias in the data — police are more likely to rearrest black

people due to racial biases. Some solutions have been proposed, such as the Algorithmic Accountability Act, which requires companies to audit their AI systems for bias in order to improve transparency and bring in public accountability. However, there are still questions remaining about whether AI will lessen or exacerbate inequities, how exactly to define fairness, and whether AI should be used for these applications at all.

Proposed Research Questions For Study: List some open-ended but driving research questions you have about the topic.	Research Questions: 1. How can the racial bias in courtroom judging AI be mitigated? 2. How can the racial bias in healthcare AI be mitigated? 3. How can the racial bias in predictive policing AI be mitigated? 4. How can the racial bias in calculating credit score AI be mitigated? 5. How can the racial bias in business be mitigated? (hiring, predictive modeling, etc.)

1.4 Draft the Problem Statement

The problem statement should be no more than 100 words and must include information covering three items (See the Actuarial Process Guide and Data Sources Guide for examples):

1. What is the risk your project is analyzing?
2. Who is at risk?
3. What are the possible risk mitigation strategies you are evaluating?

Problem statement:

Our project aims to analyze the risk derived from racial bias in decision-making AI. Typically, minorities are at risk of being misprofiled or discriminated against. Some possible risk mitigation strategies are identifying the source of error in the dataset and trying cleaning, or potentially like a counteractive filter or checker for analyzing bias in results. To do so, we are going to narrow down the scope of the project to a specific use case and analyze the AI methods in place for it.

Part 2: Data Sources

2.1 Identify Data Sources

Some data sources may have already been identified in your background research. The links found on the Modeling the Future website under Data Sources have a well-curated resource of possible data sources. As noted in The Actuarial Process Guide, other potential data sources could include government resource websites, company resources, collecting data yourself, or industry associations.

The description of your data sources should be no more than 200 words.

- List the data sources you would consider using for your project.
- Describe the value you expect each data source will bring to your project.
- What will it help you determine in your analysis, how will it help you characterize your risks, or how will it help you evaluate risk mitigation strategies?

List of Data Sources:

- https://hls.harvard.edu/wp-content/uploads/2022/08/Massachusetts-Racial-Disparity-Report-FINAL.pdf
- https://www.law.umich.edu/special/exoneration/Documents/Race%20Report%20Preview.pdf
- https://eji.org/wp-content/uploads/2005/11/race-and-the-jury-digital.pdf
- https://www.judyrecords.com/ - not that helpful
- https://www.kaggle.com/datasets/danofer/compass
- https://www.kaggle.com/datasets/usdpic/execution-database
- https://deathpenaltyinfo.org/policy-issues/innocence/exonerations-by-race

Data Sources Description:

- It provides information that quantifies the data for our topic, specifically within the state of Massachusetts.
- It provides in-depth insights into the intersection of racial bias and wrongful convictions within the criminal justice system in the United States.
- It provides background on our topic
- One of Kaggle's datasets provides information on the people who have been executed in the U.S.
- The other Kaggle data set provides information about repeat offenders
- It provides data on the death penalty exoneration rates by race

Part 3: Mathematical Modeling

3.1 Mathematical Modeling

There is no standard template for a mathematical model and it will entirely depend on the topic/project, driving research questions, data available, and assumptions made. As described in Section 3.3 of the Actuarial Process Guide, in general, your mathematical models for a Modeling the Future Challenge project should help you:

- Identify the frequencies (likelihoods) that each potential outcome for your scenario will happen.
- Identify the severities (size) of possible losses.
- Identify expected values of potential loss for your scenario.
- Understand the distribution of potential outcomes.
- Identify trends in the data and understand how potential outcomes and associated risks may be changing over time.
- Identify possible risk mitigation strategies, and quantify their effects.

Remember that your model is not expected to be perfect. You will make assumptions to create your model, and you will have uncertainties in the model results. Presenting results and acknowledging these uncertainties is important, rather than trying to say that you know exactly what will happen.

Description of your mathematical model should be no more than 200 words.

For your proposal, describe, at a high level, how you expect to mathematically model the risks for your project, and evaluate potential mitigation strategies. The description can be as simple as using regression analyses to project trends, creating a random walk model, or whatever other type of analysis you expect to conduct.

Mathematical Modeling:

The possible outcomes of our model can be broken down into true and false convictions and, within those, racially biased convictions versus those that aren't. The frequencies can be obtained once the data is broken down into a confusion matrix. Our model will fall under the paradigm of multiple regression models using the personal factors of the client and jury to output a probability of there being racial bias in the outcome of the case. We will use regression analysis on historical data about court case factors and outcomes to project trends. A simple model that can be used is a multinomial logistic regression based on the various factors in our input data mapped to the outcome. Some potential risk mitigation strategies include using this model as a supplement to inform cases regarding the potential for racial bias and ultimately promoting transparency and accountability within the judicial system. By integrating this predictive model into the decision-making process, legal professionals can be better equipped to identify and address instances of potential racial bias.

Part 4: Risk Mitigation Concepts

4.1 Risk Mitigation

In the brainstorming and research that you completed in Part 1: Project Description, you've already identified risks, parties affected, and possible risk mitigation strategies. Your problem statement was more geared toward the general topic description and risks associated. In the project phase of the MTFC, the risk mitigation strategies will inform and guide the recommendations made in the concluding sections of the project. The recommendation is "the whole point" of the project as completing work without actionable recommendations is not useful in the real world.

For the proposal, you are identifying (at a high level) what risk mitigation strategies you anticipate studying in greater detail - you will not be making recommendations yet in the proposal. This is a roadmap helping you identify where you should go with the project where you will be making recommendations.

The description of your risk mitigation strategies should be no more than 200 words.

- Describe the potential risk mitigation strategies that you think your team will evaluate for your project.
- Descriptions should be high-level to show a general concept of what possible strategies you have identified.

Given the large amount of incorrect court decisions resulting from personal biases, risk mitigation techniques are critical. COMPAS is an AI algorithm that takes in case information to predict the risk of recidivism and supposedly helps judges be less biased. However, COMPAS trains from biased past court data, only perpetuating the discrimination. To mitigate this issue, we propose an adversarial AI classifier to evaluate COMPAS's decision. This classifier will check for potential biases based on the characteristics of the defendant, such as race, gender, ethnicity, and any disability, and calculate the overall risk of a false conviction. Despite this being an added complexity within the design, its implementation in conjunction with the use of COMPAS will result in more accurate court decisions to be derived. Additionally, by addressing the bias with COMPAS, another risk mitigation technique may be utilized: implicit bias training. Given the increase in identifying racism within America, judges can undergo implicit bias training to understand their internal biases better, resulting in more accurate decisions. Having been implemented within other organizations, this risk mitigation technique proves to be useful. Consequently, by implementing both of these techniques, the amount of incorrect court case decisions will decrease in the future.