Deepa Sury

27.01.2025

Task 1.2

Is there personal information that may be exposed?

The data-set provided does not contain any PII (personally-identifiable information) and is therefore unlikely to pose any such risk to individuals or groups. However, generally-speaking, this kind of data can potentially reveal more information than intended about a particular location or its inhabitants and care must therefore be taken to ensure some amount of anonymity and confidentiality.

Are there regional or cultural biases in climate change that might be made worse by using machine learning?

The climate data is collected from 18 different weather stations over a period of more than 100 years. There are likely to be varying levels of development and technological advancement in the regions measured – this was especially true at the earliest point in time that the data was collected. Weather stations in areas with more advanced technology or higher levels of development may be overrepresented in earlier data, leading to regional bias. Using this data to make predictions would require the application of certain statistical techniques to account for this bias.

Is there human bias in climate change that might be propagated while training machine learning?

The climate data used in this project has been measured and collected by human specialists. Human bias can arise when these specialists are not adequately trained to identify certain events or patterns, or when they fail to sufficiently maintain or calibrate whatever measurement devices are used.

Additionally, historical data from several decades ago may be subject to human bias due to data collection and measuring processes being less standardized as they are nowadays.

Could machine learning potentially make incorrect decisions about where weather conditions might worsen and cause harm?

Extreme weather events have, in recent years, noticeably increased in frequency and severity across the world. When provided with the ClimateWins data-set, there is a risk of the machine-learning model downplaying these recent developments or mistaking them for noise or other irrelevant values. This can lead to the worsening of global climate problems, as countries will be less likely to take appropriate measures to reduce their environmental impact, as well as underestimation of future severe weather events, which would have devastating consequences. Care must therefore be taken to ensure that the most recent trends in climate change are accurately represented when using machine-learning to make future predictions.