ENGR 510 Ethics Homework

Anthony Su

December 6, 2024

1 Article

1.a

Price, I., Sanchez-Gonzalez, A., Alet, F. et al. Probabilistic weather forecasting with machine learning. Nature (2024). https://doi.org/10.1038/s41586-024-08252-9

1.b

Weather forecasts today rely on ensemble numerical physical simulations. An important feature of these models is their dual deterministic (realistic and detailed trajectories) and probabilistic (accurate distributions of trajectories) accuracy. The authors developed a machine learning model called GenCast which has these key features and outperforms existing numerical physical simulations in weather prediction accuracy with reduced computational cost. GenCast's architecture is based on conditional diffusion.

2 Engineering Code of Ethics

2.a

I have no experience in weather prediction and only cursory knowledge of machine learning (and none in diffusion models). I am not qualified to serve weather forecasts to the public. I would require a team of machine learning engineers (who understand diffusion models) and meteorologists (who can validate the models) in order to perform the work.

2.b

Standards which are imposed on the outputs of the simulations would be straightforward to apply. For example, data formats, standard statistical metrics, and empirical validation methods would be the same.

However, standards imposed on the model itself (such as enforcing certain conservation laws, stability measures, etc.) would be difficult to apply since these concepts may not be well-defined for the model as it is currently implemented.

2.c

The machine learning model in this case is not too unlike the model it is replacing; it is a probabilistic model. Thus, the usual caveats about uncertainty and confidence intervals apply.

2.d

Hold paramount the safety, health, and welfare of the public.

Research and development of predictive machine learning weather prediction models should always be validated to standard levels before being used to publish forecasts to the public, even at the cost of research progress. For example, if research funding to bring a model to a higher TRL is contingent on accurate published forecasts, the model must be well validated before publishing these forecasts to the public.

3 Benefits and Harms

Benefits

The immediate benefit from this model is increased accuracy. This will increase the quality of life (happiness) for all affected as plans based on weather forecasts are interrupted less frequently and dangerous weather events are forewarned more consistently.

A secondary benefit is increased equality as the barrier to entry for weather prediction is reduced. Not only is the cost of weather forecasting reduced, but also climate science education and research. These models can be used for acadeemic purposes in lower-income regions, increasing their scientific and economic autonomy.

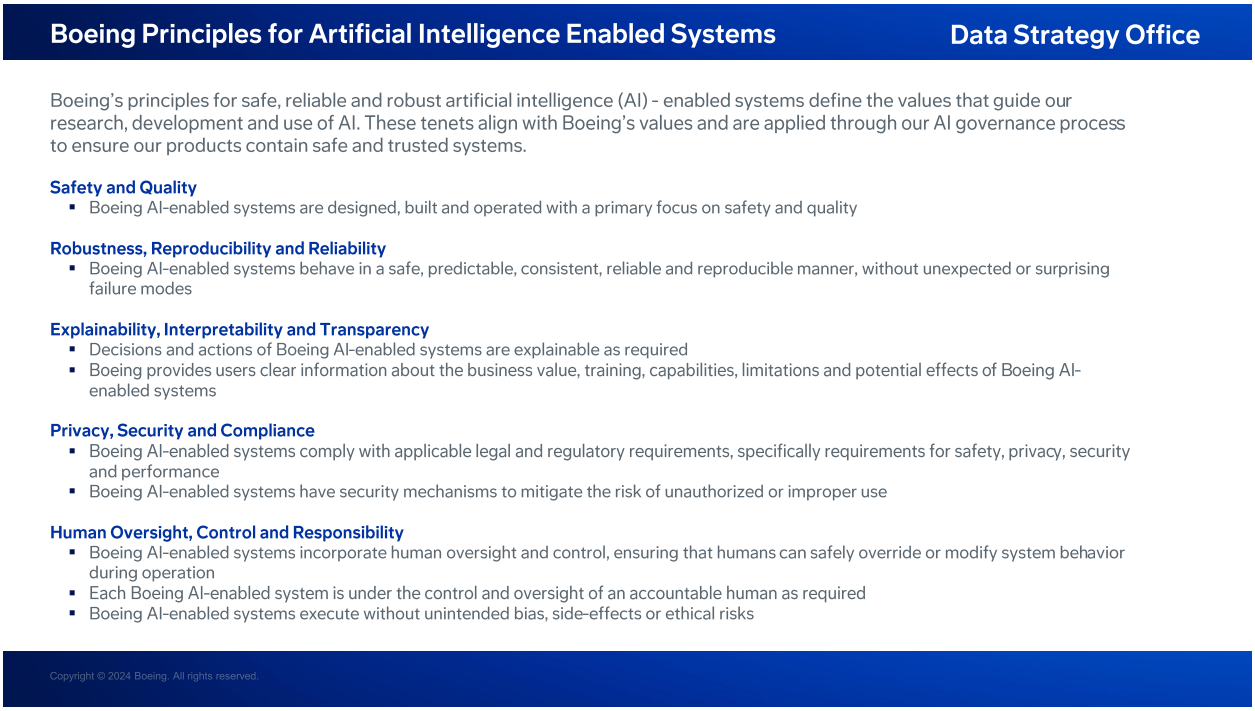
Harms

A weather model can exacerbate global inequality. If the training data for the model was not evenly sampled with uniform precision and accuracy around the globe, then wealthier regions with a higher density of weather instrumentation could be more accurately represented in the training data and thus more accurately modeled by the predictive model.

A weather model can also exacerbate global inequality by construction. GenCast utilizes an equiangular latitude-longitude grid spatial discretization. This discretization has increased grid density near the poles and decreased grid density near the equator. Thus, forecasts near the equator will be lower in resolution and potentially less accurate. This can be avoided by using a more uniformly distributed discretization such as an isocahedral grid.

4 AI or ML Code of Ethics

4.a



Boeing Principles for Artificial Intelligence Enabled Systems **Data Strategy Office**

Boeing's principles for safe, reliable and robust artificial intelligence (AI) - enabled systems define the values that guide our research, development and use of AI. These tenets align with Boeing's values and are applied through our AI governance process to ensure our products contain safe and trusted systems.

Safety and Quality
- Boeing AI-enabled systems are designed, built and operated with a primary focus on safety and quality

Robustness, Reproducibility and Reliability
- Boeing AI-enabled systems behave in a safe, predictable, consistent, reliable and reproducible manner, without unexpected or surprising failure modes

Explainability, Interpretability and Transparency
- Decisions and actions of Boeing AI-enabled systems are explainable as required
- Boeing provides users clear information about the business value, training, capabilities, limitations and potential effects of Boeing AI-enabled systems

Privacy, Security and Compliance
- Boeing AI-enabled systems comply with applicable legal and regulatory requirements, specifically requirements for safety, privacy, security and performance
- Boeing AI-enabled systems have security mechanisms to mitigate the risk of unauthorized or improper use

Human Oversight, Control and Responsibility
- Boeing AI-enabled systems incorporate human oversight and control, ensuring that humans can safely override or modify system behavior during operation
- Each Boeing AI-enabled system is under the control and oversight of an accountable human as required
- Boeing AI-enabled systems execute without unintended bias, side-effects or ethical risks

Figure 1: Boeing Principles for AI-Enabled Systems

4.b

In a few sentences, discuss what you think of the code you found. Some things you might consider: Does it seem complete to you? Does it seem relevant to the work you do? Is it easy to interpret? Is it already out of date, or do you anticipate that it will be durable, or quickly become obsolete?

The principles appear complete and relevant, especially in the context of engineered systems. However, it does not address the issue of bias which is a particularly harmful omission when data-driven models are trained on human or otherwise societal data. While this issue likely does not apply to the most obvious applications of machine learning at Boeing (e.g. system identification of engineered systems, flight anomaly detection, manufacturing quality control, etc.), Boeing products ultimately are operated by, for, and around humans and human society, and thus there is potential for unintentional reinfocement of biases.