


Predicting Weather with Machine Learning



A Presentation by Jackson Lanier

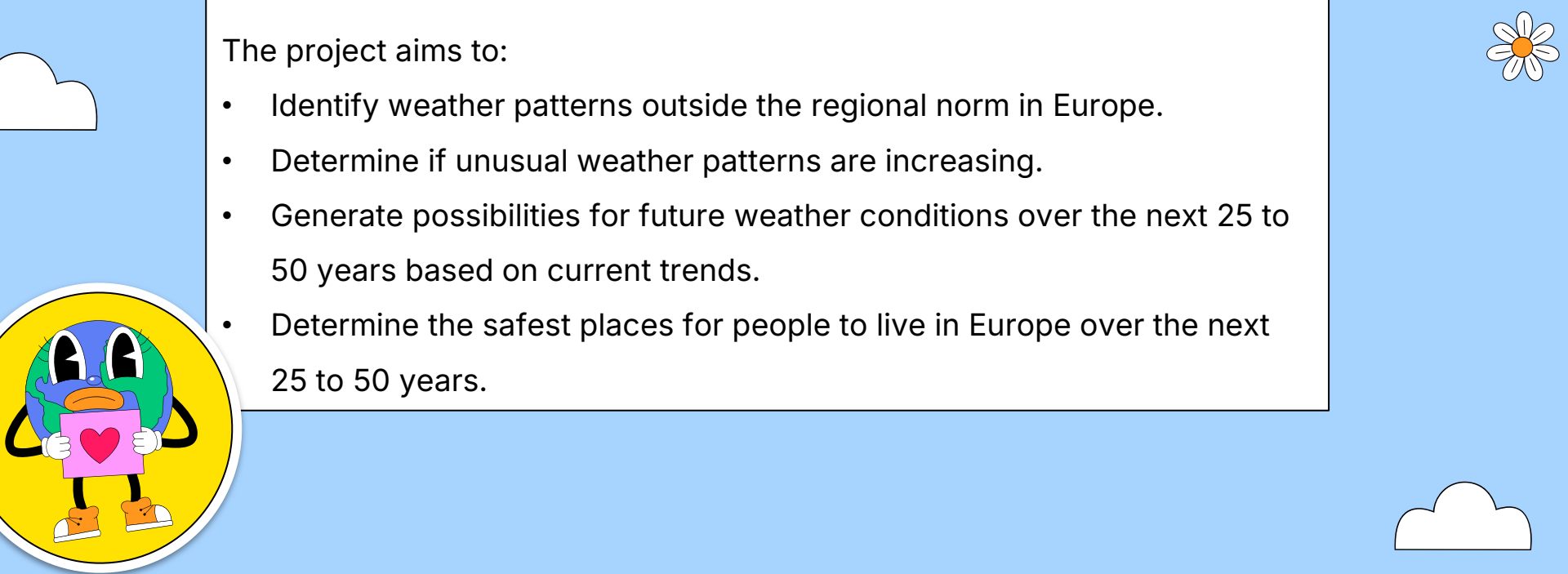
1/13/2024



ClimateWin's Objectives

ClimateWin's concern for our future has spawned a project that aims to use machine learning to help predict the consequences of climate change.

The project aims to:

- Identify weather patterns outside the regional norm in Europe.
 - Determine if unusual weather patterns are increasing.
 - Generate possibilities for future weather conditions over the next 25 to 50 years based on current trends.
 - Determine the safest places for people to live in Europe over the next 25 to 50 years.
- 

Machine Learning in Action



Thought Experiment 1

- Using random forests to assess the most impactful variables in current weather conditions
- Target various locations
- Requires current weather data

Thought Experiment 2

- Using CNNs to predict short-term weather conditions
- Focus on prediction of unusual weather events
- Requires images of weather

Thought Experiment 3

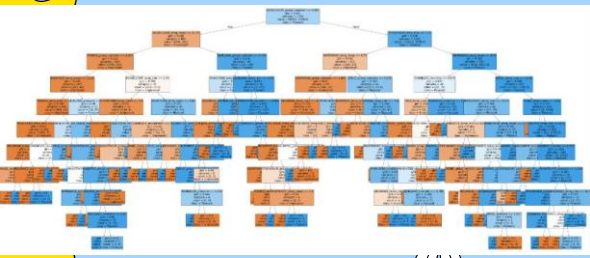
- Using GANs to create synthetic and future weather data
- Use created data to train other models
- Requires images of weather



Understanding the Models

Random Forests

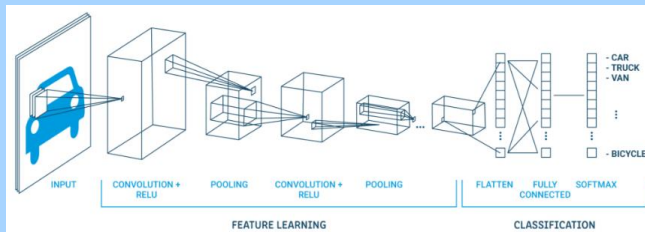
- A collection of decision trees which measure the impact of individual variables in determining an outcome
- The combined power of many trees can identify the importance of specific factors in predicting weather events



An example of a decision tree from a random forest. Imagine tens or hundreds of these trees coming together to identify impactful variables. (Image self-created)

Convolutional Neural Networks

- Drawing inspiration from the brain's visual cortex, CNNs use hidden layers to transform and weigh data, learning about the specifics of the data, such as features making up an image
- Useful for making predictions based on images such as satellite or radar images of Earth's weather

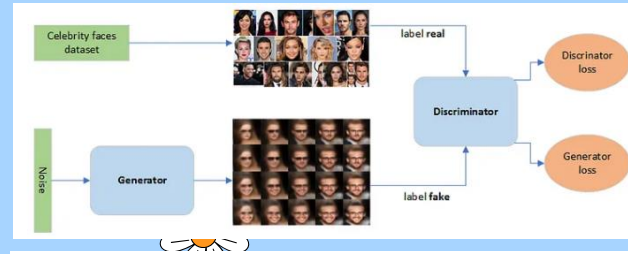


The processing method of a CNN involves convolution and pooling layers converting, compressing, and averaging pieces of data until the data reaches the output stage.

(Image source: CareerFoundry)

Generative Adversarial Networks

- GANs use competing neural networks to create, or generate, data such as images
- Models such as CNNs are used to create data while a classification model is used to determine if the generated data is fake or real
- Can be used to create satellite and/or radar images of Earth's weather patterns



The generator model creates fake data while the discriminator model tries to determine if data is real or fake. The loss measures how well each model is doing its job.

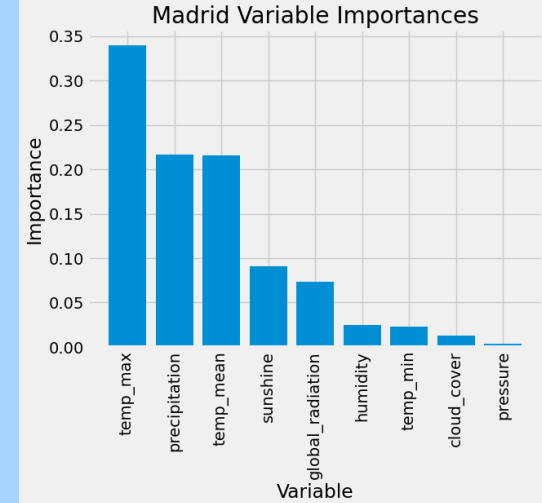
(Image source: CareerFoundry)

Thought Experiment

01

Using random forests to assess the most impactful variables in current weather conditions

- Identify specific variables, such as temperature and precipitation, that are most important in determining weather
- Focus on conditions surrounding unusual weather events, such as heat waves and flooding
- Target different locations throughout Europe
- Requires continuous, up-to-date data measuring a variety of weather conditions, similar to the data from ECA&D used throughout this project



Maximum temperature is the most important variable in determining pleasant or unpleasant days in Madrid

Potential Bias

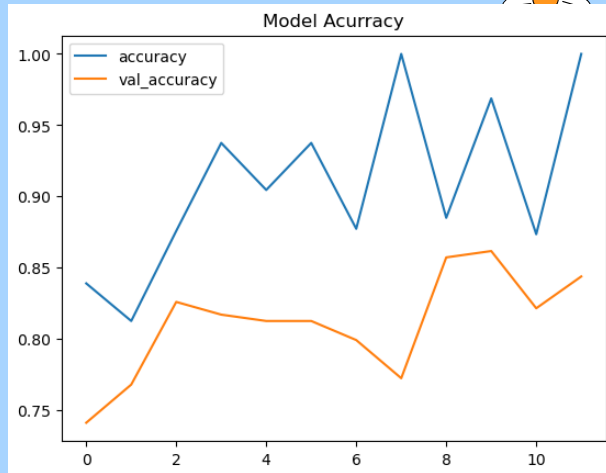
ClimateWins should pay close attention to targeted locations. For example, targeting too large an area or only a few smaller areas could lead to a bias in the results of their analyzes, leading to results that don't accurately represent impactful weather variables across Europe.

Thought Experiment

02

Using CNNs to predict short-term weather conditions

- Feed the CNN model either satellite or radar images displaying weather conditions
- Use historical images that cover a span of time, allowing the CNN to learn specific features related to different conditions
- Validation data sets should be used in testing and training the model
- Can be used to predict increases in unusual weather patterns
- Requires historical images for training and real-time images for predictions, such as radar images of historically unusual precipitation events and thermal satellite images of extreme heat events.
- Data should also include "normal" weather events, such as thermal satellite images of pleasantly warm or cool days.



During training, a model's accuracy may be higher than when exposed to new data, such as a validation data set

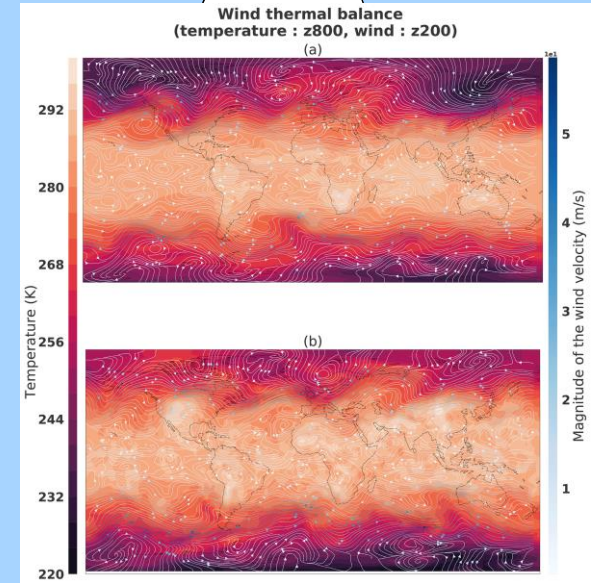
Potential Bias

ClimateWins should pay close attention to the images used in training the CNN model. Using data that is too old will underrepresent current weather conditions while not providing the model a robust amount of data could lead to inaccurate predictions on future weather.

Thought Experiment 03

Using GANs to create synthetic and future weather data

- By providing the model data on current weather events, such as satellite and radar images, GANs can generate similar images that represent future events
- Providing historical images of unusual events in order to "re-create" scenarios surrounding the events
- Provide generated data to other models for further training and analysis
- Requires both current and historical weather data, such as satellite and radar images, such as the images suggested for use in thought experiment 2




An example of comparing actual weather data to generated weather data.
[Image Source](#) (see final slide for reference)

Important!

Using GANs will require ClimateWins pay close attention to generated data and images. Although GANs can create images with a high degree of accuracy, there is the potential for the creation of misleading information, such as data that doesn't make sense (like rainstorms in the desert) or bias towards specific regions. Feeding misleading data to other models for training would cause a domino effect of inaccurate predictions.




Recommendations

- Begin by using random forests to identify impactful variables in current weather conditions, especially unusual weather events
 - Collect satellite and radar images of current weather and unusual historical weather for use in training a CNN model to understand features
 - Use GANs to generate images of future weather events and their measurements
 - Consider optimizing models for higher accuracy, such as using Bayesian optimization on the CNN model, which iteratively tests for the best hyperparameters
- 



Next Steps for ClimateWins

- Combine the power of an independent CNN model with the data generated by GANs
 - Use the random forest model to identify impactful variables in predicted and generated data
 - Assess the usefulness of each model – CNNs have the most potential
 - Use models to analyze long-term climate conditions
- 



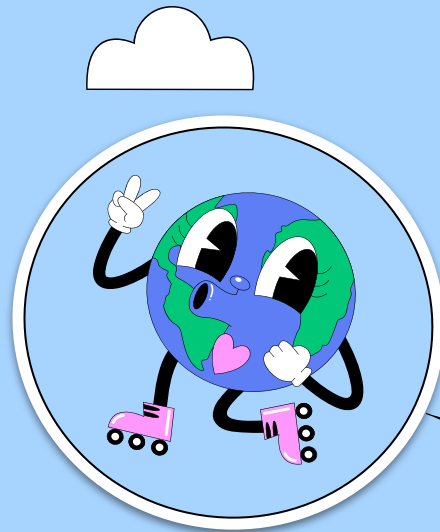
Thank You!

For questions and further discussion, please contact me at
jacksonlanier94@gmail.com

Check out my [GitHub](https://github.com/jacksonlanier) (github.com/jacksonlanier) for the Jupyter notebooks behind this work as well as additional projects!

View my video presentation on [Youtube](https://www.youtube.com/watch?v=...)!

Thermal Image in Thought Experiment 3. From *Producing realistic climate data with generative adversarial networks* by Besombes et al. (2021, July 30). Retrieved from
<https://npg.copernicus.org/articles/28/347/2021/>.



CREDITS: This presentation template was created by [Slidesgo](https://www.slidesgo.com/), including icons by [Flaticon](https://www.flaticon.com/) and infographics & images by [Freepik](https://www.freepik.com/)