Model Details

Model Name: Alberta Wildfire Prediction Model (AWPM)

Model Version: v1.0

Model Type: Classification (Random Forest for prediction, exploratory techniques for causal analysis)

Developers:

- Abdul Rahman Serhal
- Devanshi Shah
- Neda Rahimi

Release Date: Fall 2024

Primary Use: To predict wildfire risks in Alberta, including:

- **Occurrence of Fires**: Identifying high-risk areas.
- **Causes of Fires**: Classifying fires by natural or human causes.
- **Size Classifications**: Assessing the potential severity of fires.

Intended Users: Emergency response teams, insurance companies, local governments, and researchers working on wildfire prevention and management.

Out-of-Scope Use Cases:

- Predicting fires caused by extremely rare events not reflected in the training data.
- Predicting precise spread or control times due to lack of terrain data.
- Predicting wildfires outside Alberta without retraining for local data.
- Real-time dynamic predictions without continuous weather updates.

Model/Data Description

Data Used:

1. **Wildfire Historical Dataset:** Includes data on wildfire events, locations, causes, sizes, and control times.
2. **Meteorological Dataset:** Contains weather readings such as temperature, humidity, wind speed/direction, and precipitation.
3. **Additional Data Sources:** Trail data, industrial data (power, oil and gas industries), and population distribution data (for future work).

4. **Preprocessing Steps:** Missing value handling, Min-Max normalization, temporal alignment, feature engineering (rolling averages, lagged variables, seasonality).

Features:

- Meteorological indicators such as wind direction and speed.
- Temporal: Rolling averages
- Human Activity: Recreational trail and industrial data.
- Indirect environmental signals such as dryness indices (FFMC, DMC, DC).
- Data to classify **fire size**, **occurrence**, and **causes**.

Model Architecture:

  - Ensemble learning approaches including Random Forests for causal analysis and risk classification.
  - Features importance mechanism to rank predictors.
  - Consideration of weather conditions and industrial data in feature selection.

Training and Evaluation

Training Procedure:

The model was trained using historical wildfire datasets from Alberta and integrated with meteorological and activity data.

- **Random Forest Classifier**: Used for fire occurrence and cause predictions, with 100 estimators and depth optimized for generalization.
- **Additional Techniques**: Exploratory Data Analysis (EDA) informed feature selection, including weather patterns and human activity indices.

Evaluation Metrics:

- Precision, recall, and F1-score for predicting wildfire occurrences and causes.
- Performance tracked separately for different causes (e.g., natural vs. human) and size classes.

Results:

- **Occurrence Prediction**: 100% accuracy (illogical, may be due to data leakage)
- **Cause Classification**: 64% accuracy.
- **Size Classification**: Successfully classifies fires into five size categories with 94% accuracy.

Ethical Considerations

Fairness and Bias:

- The data used includes a mix of urban, rural, and remote wildfire events to minimize geographic bias. However, sparse data from remote areas may still affect predictions. Techniques like oversampling underrepresented regions help mitigate these issues.

Privacy:

- The datasets used are anonymized, and no personal or sensitive information is processed. This ensures compliance with privacy standards.

Security:

- The model should be deployed in secure environments, especially in systems handling real-time data streams, to prevent unauthorized access or misuse.

Transparency:

- Documentation on data preprocessing, feature engineering, and model performance is provided to stakeholders to ensure clear understanding and ethical use.

Limitations and Recommendations

Known Limitations:

- Sparse coverage of weather stations in remote areas affects data granularity.
- Lack of real-time terrain or slope data limits precision for fire spread prediction.
- Inability to accurately predict human-caused wildfires due to random nature.
- Difficulty modeling extreme events and rapid weather changes.

Recommendations for Use:

- Deploy as part of a broader decision-making system, not standalone.
- Regular updates with new data to address climate change impacts.
- Collaborative use with local stakeholders to refine and improve predictions

---

Additional Information

References:

1. Historical wildfire data from Alberta Forestry (eclass)
2. Meteorological data sources including local weather stations (eclass)
3. References from Alberta government websites and academic literature (included in original presentation).
4. https://www.ratehub.ca/blog/canadian-wildfires-insurance-industry/
5. https://natural-resources.canada.ca/climate-change/climate-change-impacts-forests/forest-change-indicators/cost-fire-protection/17783
6. https://wildfire.oregon.gov/evacuations
7. https://www.alltrails.com/canada/alberta
8. https://simplemaps.com/static/data/canada-cities/1.8/basic/simplemaps_canadacities_basicv1.8.zip

License: Proprietary - Available for government and insurance use under specific agreements.

Contact Information:
For further details or support, contact Devanshi Shah (ddshah1@ualberta.ca), Neda Rahimi (neda2@ualberta.ca), or Abdulrahman Serhal (serhal1@ualberta.ca)