Model Card

Model Card: Linear Regression Model for Grocery Affordability Analysis

Model Details

Model Name: Linear Regression Model for Grocery Affordability Index in Canada

Version: 1.0

Type: Linear Regression

Developer(s): Built using R's lm function, analyzing historical Canadian grocery, CPI, and wage data from multiple sources.

Date: December 2024

Input Data:

- Dataset: Data on Canadian grocery prices, CPI, and average wages from `data/02-analysis_data/grocery_data.csv`.
- Features:
- `log(CPI)`: Consumer Price Index, measuring inflation.
- `log(Average_Price)`: Average price of a representative grocery basket.
- `log(Time)`: Time variable tracking temporal changes.

Output: Predicted changes in the Affordability Index, reflecting the relationship between wages, CPI, and grocery prices over time.

Intended Use

The model analyzes grocery affordability trends in Canada by examining relationships between inflation (CPI), grocery prices, and average wages.

Intended Use Cases:

- Quantifying affordability trends over time.
- Evaluating the relationship between wages, inflation, and grocery prices.
- Asessing if wages can keep up with inflation demand.

Not Suitable For:

- Real-time grocery price forecasting.
- Predicting affordability trends outside the studied timeframe.
- Directly analyzing individual-level wage and expenditure patterns.

Performance Characteristics

Evaluation Metrics:

- R^2: 0.78, indicating strong fit to observed data.
- *RMSE*: 0.02, reflecting the model's accuracy in predicting Affordability Index trends.

Key Observations:

- Positive Relationship: log(CPI) coefficient (0.397) shows inflation correlates with slight affordability improvements, likely due to wage adjustments.
- Negative Relationship: log(Average_Price) coefficient (-0.685) demonstrates a strong negative impact of rising grocery prices on affordability.
- Gradual Improvement: log(Time) coefficient (0.024) indicates marginal improvements in affordability over time.

Considerations for Bias and Fairness

Potential Biases:

- Data reflects national averages and may not account for geographic or income-level differences.
- Data only includes a selection of groceries that may not fully caputre CPI changes.
- Data collection excludes certain population groups, such as residents in non-private dwellings or on reserves.

Fairness:

- Model does not account for regional disparities or varying household compositions, which may affect affordability differently.
- Data collection excludes certain population groups, such as residents in non-private dwellings or on reserves.

Limitations

1. **Scope of Data**: Focused only on grocery prices and wages; other cost-of-living factors are excluded.
2. **Short Timeframe**: Main analysis is limited to 2017–2023, which may not capture long-term trends.
3. **Linear Assumptions**: The model does not account for non-linear relationships or interactions between features.

Ethical Considerations

Transparency:

- Inputs, transformations, and assumptions are explicitly documented in the methodology.
- Predictions are unitless and generalized for ease of interpretation.

Responsibility:

- This model is designed for exploratory analysis and is not intended for decision-making purposes, such as wage policies or grocery pricing.

Recommendations for Future Work

1. **Data Diversification**: Expand datasets to include regional and income-level breakdowns for comprehensive analysis.
2. **Feature Consideration**: Incorporate additional cost-of-living variables, such as housing or utilities.
3. **Model Improvements**: Explore non-linear models to better capture complex interactions between variables.

Model Context

This linear regression model was built using the R programming language and the `lm` function. It explores how grocery affordability in Canada is influenced by inflation (CPI), grocery prices, and wages. It highlights the challenges posed by grocery-specific inflation outpacing wage growth, providing insights into affordability trends over time.

Contact: For questions or feedback, contact the model developer at cristina.burca@mail.utoronto.ca.