**INTRODUCTION**

Energy prices are rising, and this trend can impact customers' ability to pay their bills. Energy companies can use machine learning algorithms to predict which customers will struggle to pay their energy bills by analyzing various characteristics related to their demographic, household, and energy consumption.

This pilot study investigates the feasibility of using machine learning to predict whether a customer will struggle to pay their energy bills based on various features provided by AENERGY. It will also provide insights and recommendations to AENERGY by examining the performance of different machine learning models. The study suggests that machine learning models can predict which customers are at risk of struggling to pay their energy bills based on the available features.

**PREDICTIVE TASK**

In this scenario, the task of classifying customers into two categories—"will struggle to pay their energy bills" and "won't struggle to pay their energy bills"—must be performed. The prediction results can help AENERGY make decisions, such as creating interventions to reduce the impact of rising energy prices.

**INFORMATIVE FEATURES**

Informative features that may be good predictors for the classes include:

1. Household appliances: According to "Factors that Influence Your Bill" (n.d.), Household appliances, how many are in use, and the amount of time they are in use, are important in determining how energy efficient a home is.
2. Age of the customer: The customer’s age may indicate their ability to afford an increase in energy bills. It can suggest that a customer is of retirement age and has a fixed income, such as a pension. According to Xiao et al (2015), a positive correlation exists between age and financial capability.
3. Employment status: Unemployment correlates with financial hardship (Latsou et al., 2021). Financial hardship can cause difficulties in coping with the increase in energy bills.
4. Salary: The customer's income is directly proportional to their ability to pay their energy bills and may also have a high correlation with their credit score.
5. Credit score: A poor credit score may indicate a customer’s history of managing their finances. A person with a poor credit score will incur higher interest rates on loans and will pay more for auto, renter, and homeowner insurance (Kurt, 2022). The additional expenses can make it difficult to afford a further increase in energy bills.
6. Family size: A large household may result in higher monthly expenses, which could impact their ability to pay their energy bills (Riley, 2016).
7. Location: The customer’s residence can affect their monthly expenses and ability to pay their energy bills.

**LEARNING PROCEDURE**

The study prefers to use a decision tree algorithm for the following reasons:

* Decision trees allow for minimal data preprocessing while giving optimal results.
* Decision trees can effectively handle numerical data, such as age, salary, credit score and family size, and they can handle categorical data without problems, such as location and employment status.
* Feature scaling refers to transforming variables to have a common scale, which is often necessary for algorithms that are sensitive to the scale of input features. Decision trees and ensemble methods do not need to have the data transformed through techniques such as normalization or scaling because they are not affected by variations in the data (Thenraj, 2021).

**PERFORMANCE EVALUATION**

To evaluate the performance of the model when it is built, the dataset would be split into two sets: a training set and a validation set. The model will be trained using the training set and then used to predict labels for the observations in the validation set. The predictions will be evaluated using accuracy.

**REFERENCES**

*Factors that influence your bill*. Blue Ridge Energy. (n.d.). Retrieved January 5, 2023, from<https://www.blueridgeenergy.com/residential/help-faqs/electric/understanding-my-bill/factors-that-influence-your-bill>

Kurt, D. (2022, July 13). *The side effects of bad credit*. Investopedia. Retrieved January 7, 2023, from<https://www.investopedia.com/the-side-effects-of-bad-credit-4769783>

Latsou, D., & Geitona, M. (2021). *Effects of unemployment on financial hardship and mental health.* Psychiatriki, 32(2), 113–122.<https://doi.org/10.22365/jpsych.2021.018>

Riley, N. S. (2016, January 8). *Even family size turns out to be an inequality problem*. New York Post.<https://nypost.com/2016/01/07/even-family-size-turns-out-to-be-an-inequality-problem/>

Thenraj, P. (2021, December 15). *Do Decision Trees need Feature Scaling? - Towards Data Science*. Medium.<https://towardsdatascience.com/do-decision-trees-need-feature-scaling-97809eaa60c6>

Xiao, J. J., Chen, C., and Sun, L. (2015). *Age differences in consumer financial capability.* International Journal of Consumer Studies, 39: 387-395. Doi: 10.1111/ijcs.12205 Available at:<https://doi.org/10.1111/ijcs.12205>