1. Justificativa e importância (justificativas e importância do projeto de pesquisa em termos de relevância para a área)

The traditional power grid consists of the interconnection of several power system components including generation plants, transformers, transmission lines, substations and loads. The power flows in one direction from the generation plant to the end-consumer. The increased use of renewable energy sources such as solar, wind and geothermal in recent years has brought unique challenges to utility companies and as such the traditional power grid has been slowly shifting towards a new paradigm: the smart grid. The term “smart grid” refers to the technology that allows for two-way communication between utilities and its consumers compared to the traditional power grid where the power only flows one way. The smart grid concept is especially important when considering the increasing availability and accessibility of Internet of Things (IoT) devices and renewable energy technology. In the smart grid framework, consumers can also be energy producers and can actively engage in the energy market by planning and controlling energy usage to minimize and optimize grid stability. It is important to continue to improve the overall efficiency of power grids given that the demand for energy will only grow in the future while energy generation is costly and resources are limited.

According to Zhang et al. (2016), a key requirement for smart appliances within the smart grid framework is the demand response (DR). Power companies can provide an incentive for a consumer to actively participate in a DR program and schedule the usage of home appliances based on a dynamic price tariff. For example, customers may choose to reduce or shift electric power consumption during peak load periods. The ability to predict short and long-term home energy consumption accurately is essential to allow utility companies to provide accurate estimates of energy prices ahead of time. According to Yoon et al. (2016), home heating, ventilation and air conditioning (HVAC) are the main contributors to peak loads. The study concluded that dynamic pricing combined with dynamic thermostat controllers can result in 12% annual energy cost savings for (HVAC) operations with energy savings of up to 6% without significantly changing thermal comfort. This work analyzed the real-time price of electricity and a set of threshold prices that the thermostats used to manage energy usage during peak power price period. However, according to Zheng et al. (2016), in order to minimize power system uncertainties, energy consumers are strongly encouraged to schedule their electric energy usage based on the day-ahead price tariff. Combining the conclusions of these two papers, effective energy price prediction accompanied by the two way communication allowed by the smart grid will enable efficient day-ahead energy prediction based on weather forecast and historical consumer data which in turn can improve the long-term stability and healthy growth of the energy market.

Yu et al. (2015) has demonstrated that the distribution of meter reading can be approximated with a Gaussian distribution, which can be extrapolated to infer that any loads (e.g. deferrable loads such as HVACs) can be approximated with Gaussian distribution and estimated using stochastic methods. Deferrable loads refer to an electrical load that requires a certain amount of energy within a given time period, but the exact timing is not important. This study attempted to forecast aggregated energy usage using several machine learning based approaches such as Support Vector Machines (SVM) and neural networks.

2. Objetivos (Objetivos gerais e específicos do Projeto de Pesquisa)

Given that the literature review indicated fruitful previous attempts at predicting aggregated energy consumption for individual houses without discriminating by the load type (and load types play an important role in how the forecasting will be used), this research proposes to investigate machine learning-based techniques for energy forecasting while considering load type (fixed loads, regulatable and deferrable).

Specific objectives:

* Review literature to determine most commonly used machine learning based techniques used for energy forecasting with greater focus on SVM based approaches.
* Compare performance of different machine learning based models on energy forecast regression.
* Compare the impact of load type in performance of energy forecasting.
* Consider impact of length of training data (in days) in performance of energy forecasting.
* Apply forecasting models in real-world data (Smart\* Data Set for Sustainability).

3. Metodologia (Metodologia a ser utilizada no Projeto de Pesquisa)

Initially, a literature review will be performed around regression techniques used for energy forecasting. Real-world smart-home data from the Smart\* data set from the University of Massachusetts will be used to evaluate machine learning models and test regression performance in this experiment. The following steps will be taken with the real-world data available: 1) data pre-processing and 2) input feature selection. In addition to data processing, the input data will be grouped based on load type. Then, several machine learning-based models will be investigated using real-world data considering load type and the length of training data.

4. Cronograma (Relação itemizada das atividades previstas, em ordem sequencial e temporal, de acordo com os objetivos traçados no projeto e dentro do período de um ano)

The time-span for this project is 9 months (May 2021 to Jan. 2022). The project will be made up of the following steps in chronological order:

1. Literature review to identify state-of-the-art regression techniques used for energy forecasting. (2 months)
2. Input data processing, feature selection and load-type grouping (1 month)
3. Investigate several machine-learning based regression models using Smart\* real-world data set. (3 months)
4. Discussion of results and thesis manufacturing. (3 months)

5. Resultados e Impactos Esperados (Relação dos resultados ou produtos que se espera obter após o término da pesquisa)

The following are expected results of this work:

1. Development of an effective machine-learning based model for energy forecasting prediction.
2. Understanding how load-types affect the performance energy forecasting, particularly focused on deferrable load types such as HVACs.
3. Understanding how the length of training data (historical data of X many days prior to the requested forecasting) affects the performance of energy forecasting.
4. Implementation of the best-performing model.

6. Referências Bibliográficas (Relação itemizada das referências que subsidiam a proposta de pesquisa em ordem alfabética, com no máximo 10 referências)

Zhang, D., Li, S., Sun, M., & O'Neill, Z. (2016). An Optimal and Learning-Based Demand Response and Home Energy Management System. **IEEE Transactions on Smart Grid**, 7(4), 1790-1801.

Yu, W., An, D., Griffith, D., Yang, Q., & Xu, G. (2015). Towards statistical modeling and machine learning based energy usage forecasting in smart grid. **ACM SIGAPP Applied Computing Review**, 15(1), 6-16.

Yoon, J., Baldick, R., & Novoselac, A. (2016). Demand response control of residential HVAC loads based on dynamic electricity prices and economic analysis. **Science & Technology for the Built Environment**, 22(6), 705-719.

Smart\* Data Set for Sustainability, **University of Massachusetts Amherst**. Available at: <http://traces.cs.umass.edu/index.php/Smart/Smart>. Accessed 18 April 2021