**Determine Model accuracy scores by K-Nearest Neighbors (kNN) method:**

Effectively managing the maximum daily energy used, it is very important to analyze the energy uses in different sessions because the weather significantly affects total energy consumption. In this model, we have utilized K-Nearest Neighbors (kNN) classifier for maximum daily energy uses prediction based on classification. K-Nearest Neighbors is one of the most straightforward supervised classification techniques that provide exemplary performance results for the optimal value of K.

In this model evaluation, we have considered four stages: data processing, prediction, validation, and performance evaluation. The datasets containing weather conditions, total demand, and price categories particular period in a year that has been used in the experimentation. Two steps are involved in the prediction stage, namely the training and testing stages. In the training stage, data values are given to the classifier to train it. The training data have labels associated with them that represent their class. During the testing phase, the KNN classifier is given unlabelled data points, and the algorithm generates a list of K nearest data values. In this model, we have divided it into different training and testing features for both models to find the performance and accuracy of the predictor.

For the first model, we have trained the dataset by considering 'Minimum temperature (°C)', 'Maximum temperature (°C)', 'Rainfall (mm)', 'Evaporation (mm)', 'Sunshine (hours)' as train predictors and binning maximum energy demand in a day as a class label. We have split the train and test size into 80% and 20% in 42 random states. The highest accuracy observed for our prediction is 72.7% for daily maximum energy use based on weather conditions.

For the second model, we have trained the dataset by considering 'Minimum temperature (°C)', 'Maximum temperature (°C)', 'Sunshine (hours)', 'Evaporation (mm)', 'Speed of maximum wind gust (km/h)', '9 am Temperature (°C)' as train predictors and binning maximum daily price category in a day as a class label. We have split the train and test size into 87% and 13% in 42 random states. The highest accuracy observed for our prediction is 55.1% for the maximum daily price category based on weather conditions.

**Determine Model accuracy scores by the K-Fold method:**

K-Fold cross-validation is a method where a given dataset is split into a K number of folds. Each fold is utilized as testing set in a particular selective point, and the remaining folds train the model. This process continues until each of the folds has acted as a testing fold individually and the rest of the fold acted as a training fold. A score is retained after completing each iteration, and the sum of total iteration scores is averaged to finalize the accuracy of the model performance. We have applied the K-Fold method in both models to understand the model performance in a particular training set. First, we split the data set into ten folds and shuffled each of the folds in every test in 42 random states. We have assigned the training and testing features and class sets for each iteration by indexing (train index, test index). We have trained the model in a specific iteration using the training index of each iteration of the K-Fold process and sum up the overall class test and prediction value.

For the first model, we have trained the feature dataset. The model accuracy observed for our prediction is 49.9% for daily maximum energy use based on weather conditions.

For the second model, we have trained the feature dataset. The model accuracy observed for our prediction is 49.9% for daily maximum energy use based on weather conditions.

**Discussion and conclusion**

The purpose of building these models is to predict the maximum daily energy used and energy pricing based on weather conditions that significantly help the energy companies to understand how weather conditions may influence the overall energy demands and the effect of several variables on energy prices that can help them a detailed understanding of their future business model based on future energy uses.

For both models, we used supervised algorithms. The model between the maximum daily energy consumption and average temperature, due to MAXDEMAND is continuous data, Regression analysis is suggested for modelling and finding the cause and effect relationship between variables. For the model which predicts the maximum daily price category based on the weather data, since price category is labelled, KNN is a good option to build a model for discrete variables.