**IT4060 – Machine Learning**

**Lab 7 – Random Forest**

**Random Forest Classifier**

1. Upload the *Random Forest Digits* python notebook to Jupyter notebook and run it. It uses the Random Forest Classifier to classify the images of digits.

**Note:** You may have to change the path of the input data file.

TODO: The example uses the classification\_reports function in the metrics module. Refer the scikit-learn documentation and explain the output given by this function.

TODO: Then, the confusion matrix prints a matrix using he confusion matrix module. Refer the scikit-learn documentation and explain the meaning of the confusion matrix.

**Random Forest Regression**

1. Upload the *Random-forest-regression-time-series-prediction.ipynb* file to Jupyter notebook and run it. It predicts the temperature of a given day based on a historical time series of temperature values. It uses the Random forest regression classifier. It is based on the following online tutorial.

**Note:** You may have to change the path of the input data file.

<https://towardsdatascience.com/random-forest-in-python-24d0893d51c0>

TODO: The tutorial contains code snippets to consider only the most significant features to do the regression analysis. It is done using the *feature\_importance\_* property of the random forest regression model. Add those snippets to do the regression analysis again, only considering the most relevant input features.

**Submission:**

Upload the html files exported by Jupyter notebooks as a single zip file to the courseweb link. Also, add a separate text file containing the answers for the first exercise. The combined zip file name should be your registration number.