1.Business and Data Understanding

“O **training set** deverá ser usado para desenvolver e treinar o modelo de Machine Learning. No training set é fornecida informação referente ao nível de incidentes rodoviários (**accidents**) de cada registo para que possam treinar os modelos de aprendizagem. O modelo a desenvolver deverá ter, na sua base, features como a magnitude do atraso que se verifica numa determinada hora, o tempo de atraso provocado pelos incidentes, a temperatura, pressão atmosférica e a velocidade do vento, entre outras features que caracterizam um determinado ponto temporal.”

“O **test set** será utilizado para validar a performance do modelo em dados ainda não vistos pelo mesmo. Neste dataset de teste não é fornecida a class referente ao nível de incidentes rodoviários. O vosso trabalho é prever essa mesma class. Devem utilizar o modelo desenvolvido para prever, para cada registo do test set, o nível de incidentes rodoviários esperados.”

É também disponibilizado um exemplo de um ficheiro de submissão (example\_submission.csv) onde é assumido que o nível de incidentes rodoviários esperado para cada registo é Medium.