Laborationsrapport 2~ DV2621: Artificiell Intelligens ~  
Datum: 12-12-2022  
Namn gruppmedlem 1: Suvapit Kassemchaikulapat  
E-post gruppmedlem 1: suks17@student.bth.se  
Namn gruppmedlem 2: Engla Nilsson  
E-post gruppmedlem 2: enni22@student.bth.se  
Del 1  
Uppgift 1.1  
Uppmätt accuracy: 0,70  
Uppgift 1.2  
Uppmätt accuracy: 0,70  
Uppgift 1.3  
Uppmätt accuracy: 0,83   
Uppgift 1.4  
Förklara vad tror ni de stora skillnaderna i accuracy beror på i de tre uppgifterna ovan?  
Svar: Vi tror att skillnaden beror på storleken av dataframen dvs att en större dataframe ger högre accuracy då det är fler attribut inom dataframen som väger upp accuracyn.  
Del 2  
Uppgift 2.1  
Antal instanser: 712  
Antal attribut per instans: 10  
Uppgift 2.2  
Vilka olika typer av attribut innehåller datan, och hur skiljer de olika typerna sig åt: int, object, float. Int och float är siffror medans objekt är icke numerisk data. Int är bara heltal, float är decimaltal.   
Uppgift 2.3  
Antal Female=... (varav 74,2% omkom)  
Antal Male=... (varav 18,9% omkom)  
Uppgift 2.4  
Klass-ratio: 6:10  
Är datamängden balanserad: nej

Del 3  
Uppgift 3.1  
DecisionTree-accuracy: runt 71%  
Uppgift 3.2  
RandomForest-accuracy: 75%  
Uppgift 3.3  
Vad är er slutsats till varför accuracymåtten skiljer sig åt: Vår slutsats är att accuracymåtten skiljer sig åt för att descition tree och randomforrest klassifikationer är byggda på olika sätt och ger därför olika accuracy.  
Frivilliga extrauppgifter  
Uppgift 3.4  
Svar: Nej det är det inte för att det inte är stabilt. Ett bra mätvärde är precision istället för accuracy är att ändra på split data. Dock hittar vi inget bättre score än accuracy. Av alla score så blir accuracy alltid högst på randomforrest och decisiontree

Uppgift 3.5  
Svar: tar man då splitdata till 50 % blir accuracyn då istället 78%  
Uppgift 3.6  
Svar: Vi hittar ingen hyperparameter på internet som ger högre procent än accuracy.