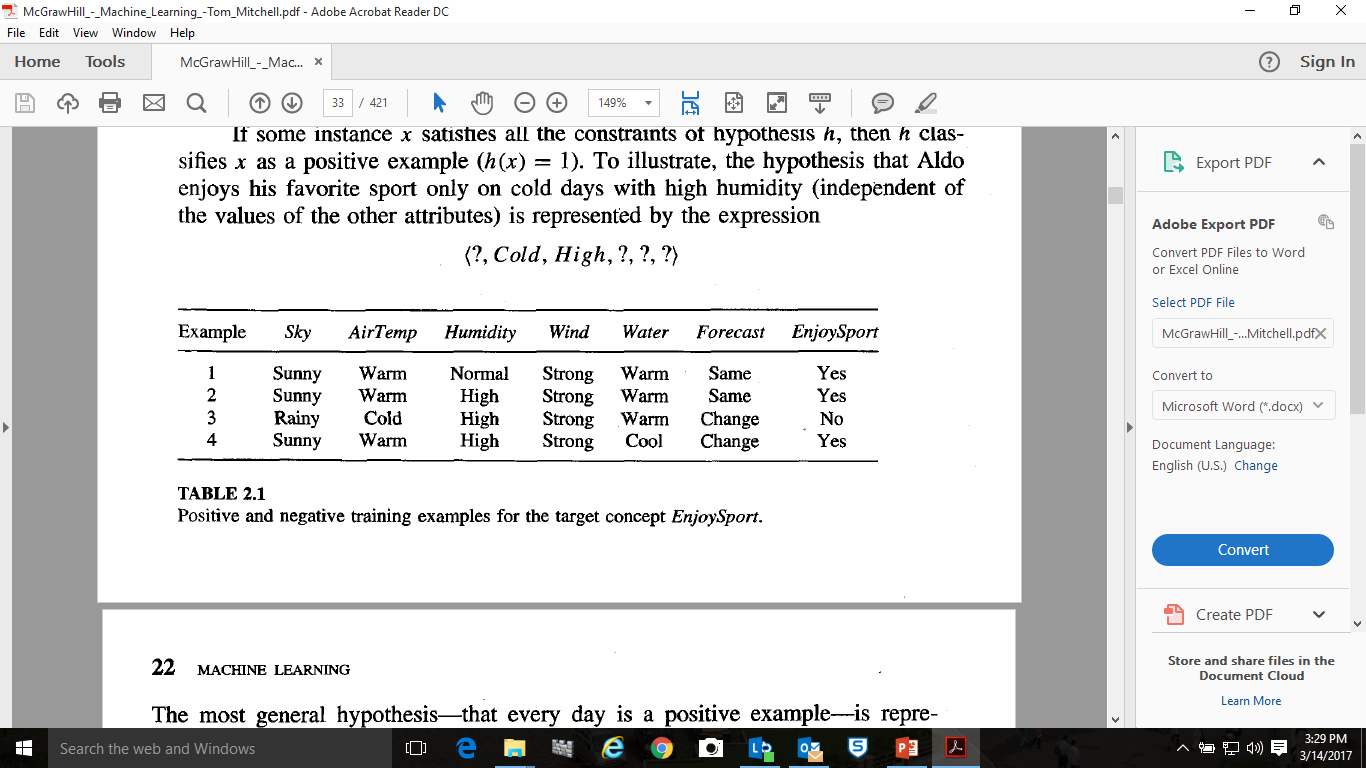
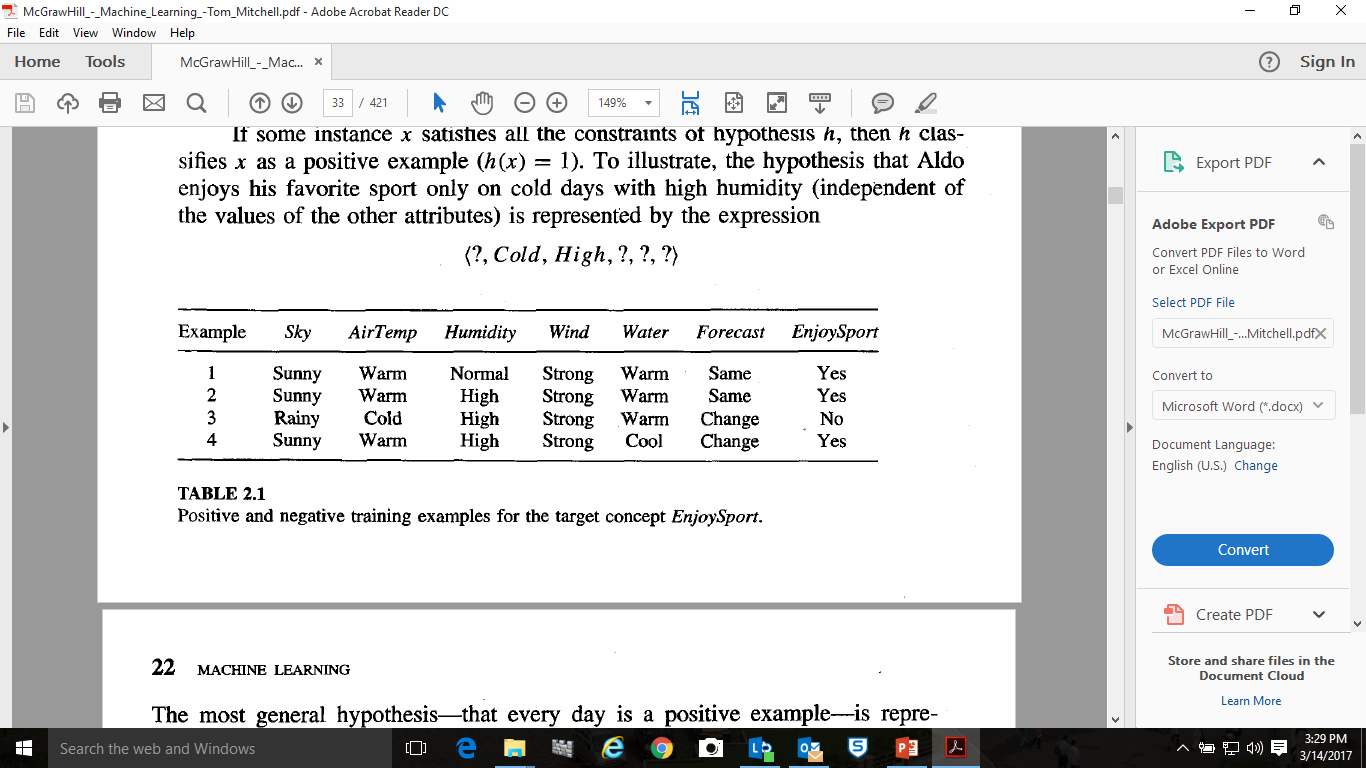
BDA 601L Fundamentals of Machine Learning Lab

Assignment

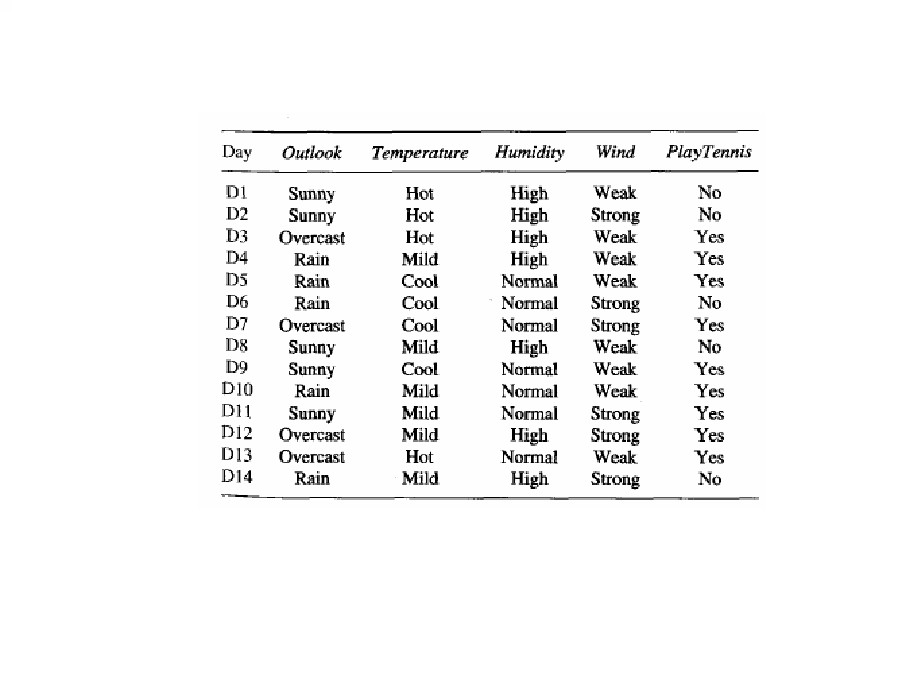
1. Obtain the hypothesis space search by Find-S for the given training examples.



1. Apply Candidate Elimination Algorithm for the given training examples to obtain a hypothesis space.



1. Use ID3 Algorithm to develop the decision tree support system for the following training data set.



1. Data set given in the table is for a company produce tissues (used by biological labs). Company objective is to predict how well their products are accepted by the clients. They conducted a survey with their clients to find the acceptance levels of the products. As shown in figure, Type- 1 and 2 are not well accepted whereas Type- 3 and 4 are well accepted.

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Acid Durability** | **Strength** | **Class** |
| Type-1 | 7 | 7 | Bad |
| Type-2 | 7 | 4 | Bad |
| Type-3 | 3 | 4 | Good |
| Type-4 | 1 | 4 | Good |

Built a classifier to find the class for the following new type with K=3:

**New type: Type-5; Acid Durability = 3; Strength = 7**

1. Apply the naive Bayes classifier to a concept learning problem to classify days according to whether someone will play tennis. The table given below provides a set of 14 training examples of the target concept **PlayTennis**, where each day is described by the attributes ***Outlook, Temperature, Humidity, and Wind***.



Use naive Bayes classifier and the training data from this table to predict the target value (***yes*** or ***no***) of the target concept ***PlayTennis*** for the following new instance:

***(Outlook = sunny, Temperature = cool, Humidity = high, Wind = strong)***

***(Use the data set already shared with you)***

1. ***Develop a Decision Tree for the given data set and find the accuracy of the model.***
2. ***Develop a Naïve Bayes Classifier for the given data set and find the accuracy of the model.***
3. ***Develop a Logistic Regression classifier for the given data set and find the accuracy of the model.***
4. ***Develop a K-NN model for the given data set and find the accuracy of the model.***
5. ***Develop a linear regression for the given data set and find the accuracy of the model.***
6. ***Develop a Logistic Regression model and apply PAC for the given data set and find the accuracy of the model.***
7. ***Develop a random forest for the given data set and find the accuracy of the model.***
8. ***Compare the performance of all the classifiers***
9. ***Write the Python code to compute entropy and information gain***
10. ***Write the Python code to demonstrate conditional probability***
11. ***Write the Python code to compute Euclidean Distance between data points***
12. ***Write the Python code to calculate covariance matrix, Eigen values and Eigen vectors***
13. ***Write the Python code to calculate the following***
    1. ***Accuracy e***
    2. ***Misclassification***
    3. ***Type-1 and Type-2 error rates***
    4. ***Sensitivity***
    5. ***Specificity***