Exercise 1

1. Alice wants to write an application to determine if a text is discussing about scientific policy based on the frequency of the words *science*, *public*, *access*, *university*, *government*, *finance*, *education*, *budget*, *justice*, and *law*. She labeled millions of texts according to their subject. What kind of automatic learning problem she must solve?
2. Are the following problems machine learning problems? If yes, of what type?
    1. Given a car's owner manual and the price of petrol at the nearest filling station, predict how much will cost to fill up the car's tank?
    2. Determine the optimal time to publish a text in a given Web platform.
    3. Compute the shortest path between two nodes in a graph
    4. Remove the noise of a radio signal
    5. Predict the number of bicycles to make available in a bike rental station
    6. Compute a house's heating load (i.e., the amount of energy that is needed to maintain the temperature) from its building plan and a civil engineer's records of plans and heatings loads for houses in the same neighborhood.
3. Bob's data set comprises 10 variables. He would like to visualize using two-dimensional representation. What type of algorithm he must use?
4. Eva wants to examine her spams to check if there are subtypes of spams. What learning algorithm she must use?
5. We have seen that Tom Michell [] defines machine learning as: "A computer program is considered to learn from experience **E** with respect to some class of tasks **T** and performance measure **P** if its performance tasks in **T**, as measured by **P**, improves with experience **E**". Alice is developing an application that uses the historical financial transactions to detect frauds. In this case, what are the **E**, **T**, and **P**?
6. What is meant by the term inductive bias?
7. Why is machine learning considered an ill-posed problem?