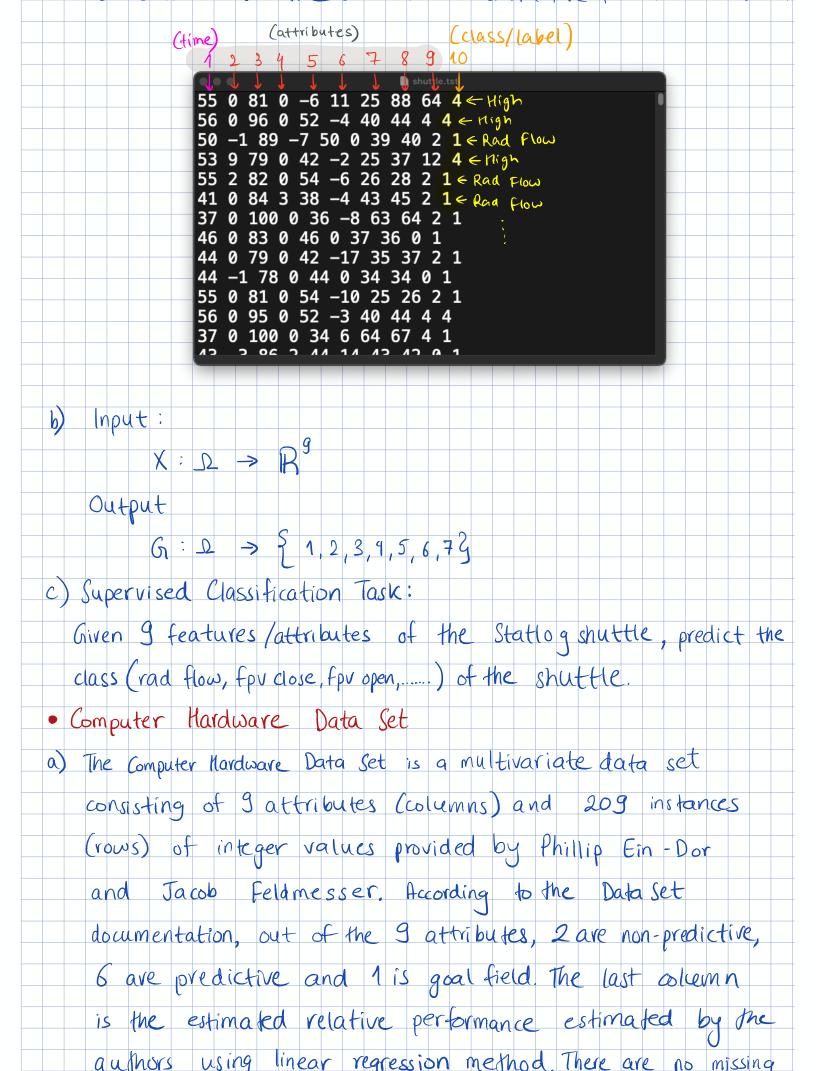
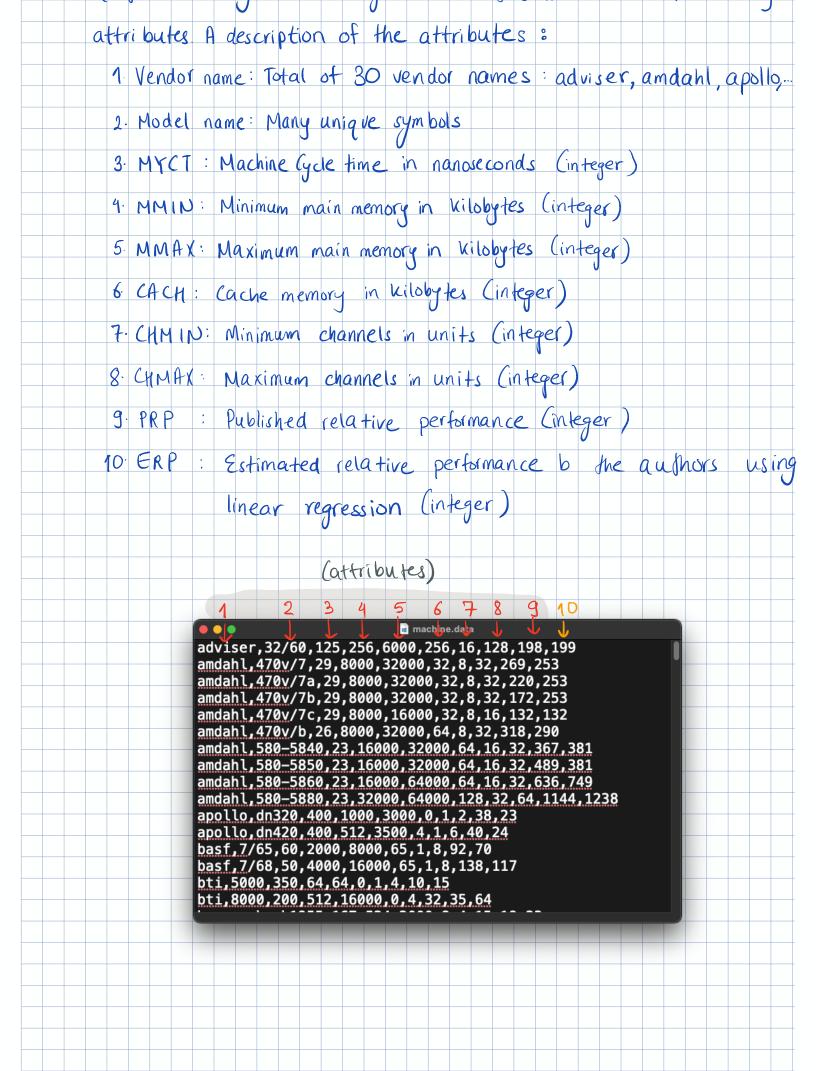
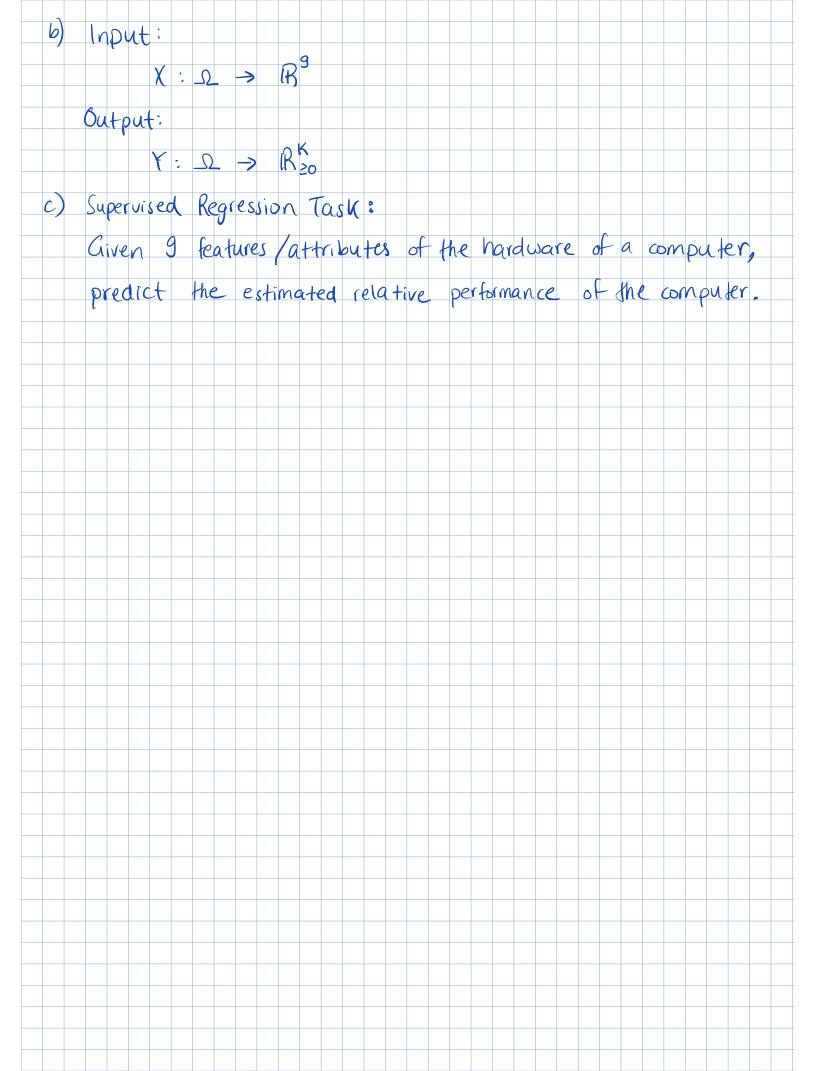
## ASSIGNMENT 3 SOLUTION Done By: Nayan Man Singh Pradhan Excercise 1. (Modelling inputs / outputs) In this exercise you work with two data sets: • Statlog (Shuttle) Data Set • Computer Hardware Data Set, which are both available in the UCI Machine Learning Repository. For each of them, perform the following tasks: a) Briefly describe the data set and all involved variables in your own words. If some information is missing on the UCI Repository site, do your own search for these details. b) Model the data set via input and output random variables / vectors. c) Formulate a question that can be solved using machine learning on this data set and give the type of machine learning (supervises / unsupervised / regression / classification) that will allow to answer the question. (8 Points) Station (Shuttle) Data Set The Statlog (Snyttle) Data Set is a multivariate data set consisting of 9 attributes (columns) and 58,000 instances (rows) of integer values provided by Tason Catlett from NASA'S shuttle datasets. According to the Data Set documentation, the shuttle dataset contains 9 attributes, all of which are numerical values. The first attribute is "time". and the last column is the class / label. There are total 7 classes / labels They are: 1 Rad Flow 5 Bypass 6. Bpv Close 2. For Close 3. For Open Bpv Open 4. High The documentation specifies that approximately 80% of the data belongs to class 1. I have attached a labelled screensnot of the downloaded data file for clarification







Excercise 2. (SPAM e-mail representation)  The Spambase Data Set is a SPAM classification data set that has exactly the 57 input variables that are roughly described in Example 2.9 of the lecture.																																				
a) In Example 2.9, we did not mention the specific words and characters that are used in the features. information of the UCI Repository to make a complete description of these variables, i.e. give all the key etc.																																				
	b)	Sea	rch					alter and									e us	sed 1	to c	lesc	ribe	e (S	SPAN	<b>I</b> ) e	ma	ils. P	ick	one	exa	ample feature —						
		set,	, CI	.е т.	ne s	oui	ce,	anc	ı de	SCI	ње	une	se i	eau	ures	٠.										(4 Po										