

Hands on

Machine Learning on gene expression dataset

Gene Expression Dataset (RNASeq)

uterus_rnaseq_split_1.tsv - LibreOffice Calc

File Edit View Insert Format Styles Sheet Data Tools Window Help

Liberation Sa 10 B I U A

A1 fx Σ =

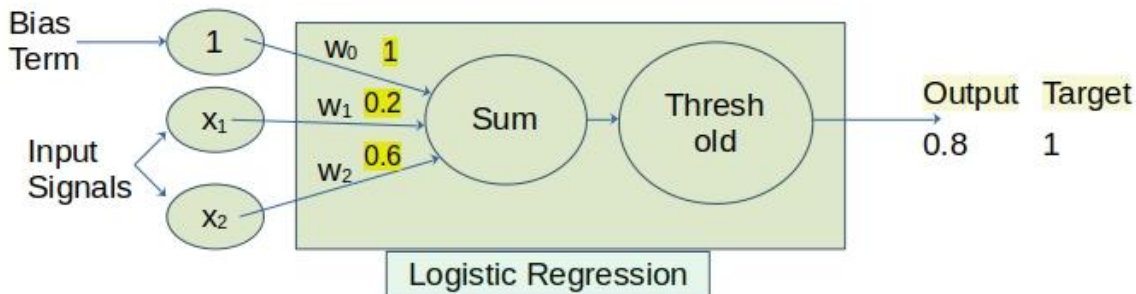
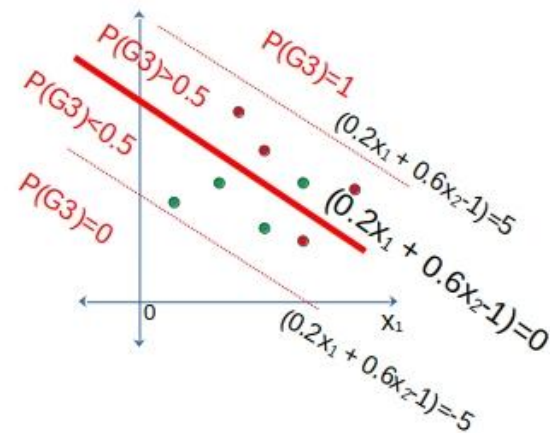
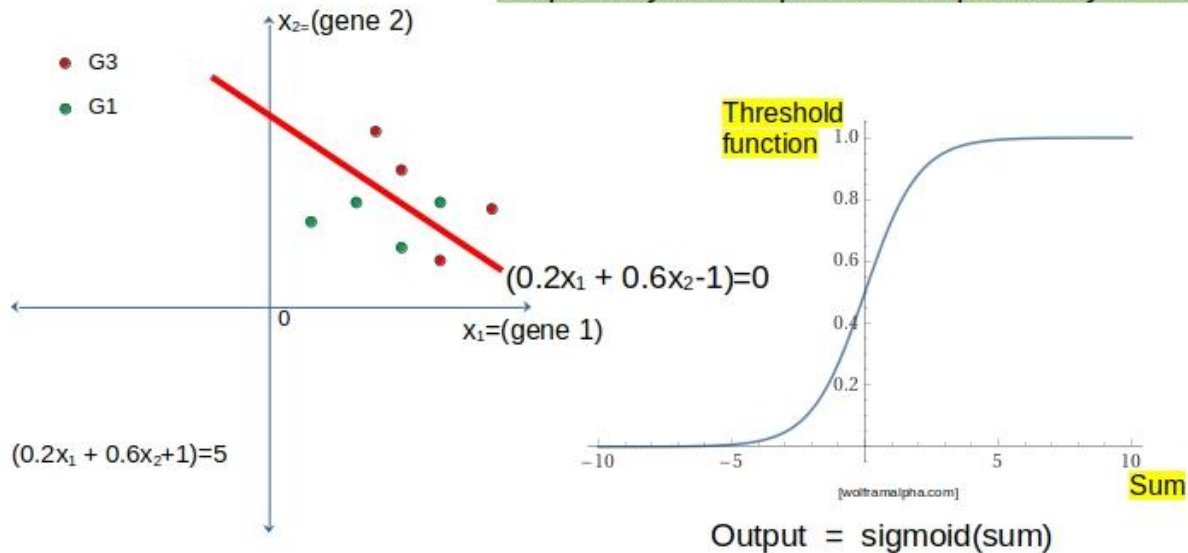
	A	B	C	D	E	F
1		ENSG000000000003	ENSG000000000005	ENSG000000000419	ENSG000000000457	ENSG000000000000
2	TCGA.4E.A92E.01	11.1000658477563	3.48567278072098	9.23704848979341	9.50621601002227	8.13317683365
3	TCGA.A5.A0G9.01	10.9655038437165	4.1821	11332859199	9.00317111330623	7.46077790970
4	TCGA.A5.A0GA.01	10.6155212878745	2.8223	17977273491	9.77050409028395	8.94614994156
5	TCGA.A5.A0GB.01	11.6419860817083	2.82239645514985	9.56374621528024	8.66988447495112	8.87794170134
6	TCGA.A5.A0GG.01	11.2623545882775	3.56267484979439	9.14005719428663	8.7952188286193	8.08187091367
7	TCGA.A5.A0GH.01	11.5658458457742	3.70329059933182	9.4780154849539	8.86002914690886	9.13596425384
8	TCGA.A5.A0GM.01	11.8698930989358	2.82239645514985	8.8098773076786	8.23578944925482	6.64428673960
9	TCGA.A5.A0GN.01	10.1783570137085	2.82239645514985	7.97672937468917	10.5579336299739	7.49207399479
10	TCGA.A5.A0GP.01	11.7123135859621	3.61685538796476	9.87678524159442	8.45804227083872	7.95544727564
11	TCGA.A5.A0GQ.01	11.8284540125815	4.0051354988623	9.24294711747796	8.47997399750973	7.14937714371

Feature

Feature value

Logistic Regression -> Linear classifier.

Output may be interpreted as a probability value for the class having label 1.



Instructions

1. Download the zip file from the Google Drive Link provided.
Day3_gene_hemant.zip
2. Unzip the zipped file in your PC/Laptop.
Day3_gene_hemant.zip -> Day3_gene_hemant
3. Upload the folder into you Google Drive.
Day3_gene_hemant -> Day3_gene_hemant
4. Double click to go inside the folder in Google Drive.
Day3_gene_hemant
5. Click on the part1_day3_9.45am file and open it with google colab.
part1_day3_9.45am