Assignment #2

Elements of Machine Learning

Saarland University – Winter Semester 2024/25

Rabin Adhikari
7072310
raad00002@stud.uni-saarland.de

Dhimitrios Duka
7059153
dhdu00001@stud.uni-saarland.de

Problem 2 (Logistic Regression)

a)

In order to derive the optimal coeffiecients, we will have to take the derivative of the loss function $l(x)$ w.r.t each β_j. First, we will derive a generic expression for camputing the gradient w.r.t β_j.

$$\begin{aligned}
\frac{\partial}{\partial \beta_j}\ell(\beta) &= \frac{\partial}{\partial \beta_j}\sum_{i=1}^{n}\left[y_i \log p(x_i;\beta) + (1-y_i)\log\left(1-p(x_i;\beta)\right)\right] \\
&= \sum_{i=1}^{n}\left[y_i \frac{\partial}{\partial \beta_j}\log p(x_i;\beta) + (1-y_i)\frac{\partial}{\partial \beta_j}\log\left(1-p(x_i;\beta)\right)\right] \\
&= \sum_{i=1}^{n}\left[y_i \frac{\frac{\partial}{\partial \beta_j}p(x_i;\beta)}{p(x_i;\beta)} + (1-y_i)\frac{\frac{\partial}{\partial \beta_j}(1-p(x_i;\beta))}{(1-p(x_i;\beta))}\right] \\
&= \sum_{i=1}^{n}\left[y_i \frac{\frac{\partial}{\partial \beta_j}p(x_i;\beta)}{p(x_i;\beta)} - (1-y_i)\frac{\frac{\partial}{\partial \beta_j}p(x_i;\beta)}{(1-p(x_i;\beta))}\right] \\
&= 0
\end{aligned} \tag{1}$$

b)

c) i)

Applying the formula for a multivariate logistic regression, we get the following outputs:

Table 1: Data for x1, x2, and predicted values.

x1	**x2**	$p(x_i, \beta)$	**class**	**correct**
1.0	2.0	0.182	0	0
2.0	3.0	0.378	0	0
3.0	4.0	0.622	1	0
4.0	5.0	0.818	1	1
5.0	6.0	0.924	1	1
6.0	7.0	0.971	1	1
7.0	8.0	0.989	1	1
8.0	9.0	0.996	1	1

c) i)

Based on the given theshold, we can construct the following table. we see that the model missclassified only one point.