Artificial Intelligence for Sciences and Humanities

Homework 6: The Cornerstone of AI: Gradient Descent

DO NOT POLLUTE! AVOID PRINTING, OR PRINT 2-SIDED MULTIPAGE.

In this homework you will implement gradient descent and use it to construct a logistic regression AI system.

To this end, we will use the titanic dataset (`titanic_data.csv`), containing the following information about 887 passengers: 1) whether they survived or not (1 = survived, 0 = deceased), 2) passenger class, 3) gender (0 = male, 1 = female), 4) age, 5) number of siblings/spouses aboard, 6) number of parents/children aboard, and 7) fare:

	Passenger 1	Passenger 2	Passenger 3	···	Passenger 887
Survived	**0**	**1**	**1**	···	**0**
Passenger Class	3	1	3	···	3
Gender	0	1	1	···	0
Age	22	38	26	···	32
Siblings/Spouses	1	1	0	···	0
Parents/Children	0	0	0	···	0
Fare	7.25	71.2833	7.925	···	7.75

Our goal is to construct an AI system that determines/predicts whether an individual would survive or not.

Each problem is worth 50 points.

Problem 1. Deliver python code to load this dataset into data vectors $\mathbf{x}_i$ and their corresponding responses y_i.

Problem 2. Recall that logistic regression uses the following log-likelihood as a loss function:

$$\ell(\mathbf{w}) \;:=\; -\sum_{i=1}^{N} \log\left[\left(\frac{1}{1+e^{-\mathbf{w}^\top \mathbf{x}_i}}\right)^{y_i}\left(\frac{1}{1+e^{\mathbf{w}^\top \mathbf{x}_i}}\right)^{1-y_i}\right],$$

whose gradient is given by:

$$\ell'(\mathbf{w}) \;=\; -\sum_{i=1}^{N}\left(y_i - \frac{1}{1+e^{-\mathbf{w}^\top \mathbf{x}_i}}\right)\mathbf{x}_i.$$

Deliver python code to compute this gradient at a given $\mathbf{w}$.

Problem 3. Recall that gradient descent can be summarized as an iterative process that repeats until convergence the following update:

$$\mathbf{w}_{t+1} \;=\; \mathbf{w}_t - \eta\, \ell'(\mathbf{w}_t),$$

where $\mathbf{w}_0$ is an initial *guess*, and η is the step size parameter. Deliver python code to implement gradient descent.

Problem 4. Use your implementations above to learn the optimal parameter $\mathbf{w}^\star$. What value do you obtain for $\mathbf{w}^\star$?

Problem 5. What value or strategy did you use for η?

Problem 6. Change η to another value and repeat the training process. What value did you obtain for $\mathbf{w}^\star$?

Problem 7. Explain which choice of η is better, and why.

Problem 8. Build your own feature vector $\mathbf{x}$. According to your AI system, would you have survived the titanic sinking? Briefly explain your conclusion.