

```

1 import numpy as np
2 import util
3 import matplotlib.pyplot as plt
4
5 def main(lr, train_path, eval_path, save_path):
6     """Problem: Poisson regression with gradient ascent.
7
8     Args:
9         lr: Learning rate for gradient ascent.
10        train_path: Path to CSV file containing dataset for training.
11        eval_path: Path to CSV file containing dataset for evaluation.
12        save_path: Path to save predictions.
13    """
14    # Load training set
15    x_train, y_train = util.load_dataset(train_path, add_intercept=True)
16
17    # *** START CODE HERE ***
18    # Fit a Poisson Regression model
19    clf = PoissonRegression(step_size=lr)
20    clf.fit(x_train, y_train)
21
22    # Run on the validation set, and use np.savetxt to save outputs to save_path
23    x_eval, y_eval = util.load_dataset(eval_path, add_intercept=True)
24    p_eval = clf.predict(x_eval)
25    np.savetxt(save_path, p_eval)
26    plt.figure()
27    plt.scatter(y_eval, p_eval, alpha=0.4, c='red', label='Ground Truth vs Predicted')
28    plt.xlabel('Ground Truth')
29    plt.ylabel('Predictions')
30    plt.legend()
31    plt.savefig('poisson_valid.png')
32    # *** END CODE HERE ***
33
34
35 class PoissonRegression:
36     """Poisson Regression.
37
38     Example usage:
39     > clf = PoissonRegression(step_size=lr)
40     > clf.fit(x_train, y_train)
41     > clf.predict(x_eval)
42     """
43
44     def __init__(self, step_size=1e-5, max_iter=10000000, eps=1e-5,
45                 theta_0=None, verbose=True):
46         """
47         Args:
48             step_size: Step size for iterative solvers only.
49             max_iter: Maximum number of iterations for the solver.
50             eps: Threshold for determining convergence.
51             theta_0: Initial guess for theta. If None, use the zero vector.
52             verbose: Print loss values during training.
53         """
54         self.theta = theta_0
55         self.step_size = step_size
56         self.max_iter = max_iter
57         self.eps = eps
58         self.verbose = verbose
59
60     def fit(self, x, y):
61         """Run gradient ascent to maximize likelihood for Poisson regression.
62
63         Args:
64             x: Training example inputs. Shape (n_examples, dim).
65             y: Training example labels. Shape (n_examples,).
66         """
67         # *** START CODE HERE ***
68         m, n = x.shape
69         if self.theta is None:
70             self.theta = np.zeros(n, dtype=np.float32)
71
72         prev_theta = None

```

```

73     i = 0
74     while i < self.max_iter \
75         and (prev_theta is None
76             or np.sum(np.abs(self.theta - prev_theta)) > self.eps):
77         i += 1
78         prev_theta = np.copy(self.theta)
79         self._step(x, y)
80         if self.verbose and i % 5 == 0:
81             print('[iter: {:02d}, theta: {}]'
82                   .format(i, [round(t, 5) for t in self.theta]))
83         # *** END CODE HERE ***
84
85     def predict(self, x):
86         """Make a prediction given inputs x.
87
88         Args:
89             x: Inputs of shape (n_examples, dim).
90
91         Returns:
92             Floating-point prediction for each input, shape (n_examples,).
93         """
94         # *** START CODE HERE ***
95         y_hat = np.exp(x.dot(self.theta))
96
97         return y_hat
98
99     def _step(self, x, y):
100         """Perform a single gradient ascent update step."""
101         grad = np.expand_dims(y - np.exp(x.dot(self.theta)), 1) * x
102         self.theta = self.theta + self.step_size * np.sum(grad, axis=0)
103         # *** END CODE HERE ***
104
105 if __name__ == '__main__':
106     main(lr=1e-5,
107         train_path='train.csv',
108         eval_path='valid.csv',
109         save_path='poisson_pred.txt')

```