CS189 Homework 7: IM2SPAIN

May 2024

1 Overview

In this homework, k-nearest neighbors (k-NN) was utilized to estimate geographic locations (longitude and latitude) from a given query photograph. The dataset comprises geo-tagged images from Flickr, focusing on images within Spain, contrasting with the extensive dataset used in the original IM2GPS by Hays and Efros [1], which leveraged over 6 million photographs from the internet, including Flickr. Each image was processed with OpenAI's CLIP image model to produce features for use in k-NN.

2 Plotting Features

Using matplotlib and scikit-learn, the data was visualized on a graph, plotting the geographic location of each image, with the x-axis representing longitude and the y-axis representing latitude.

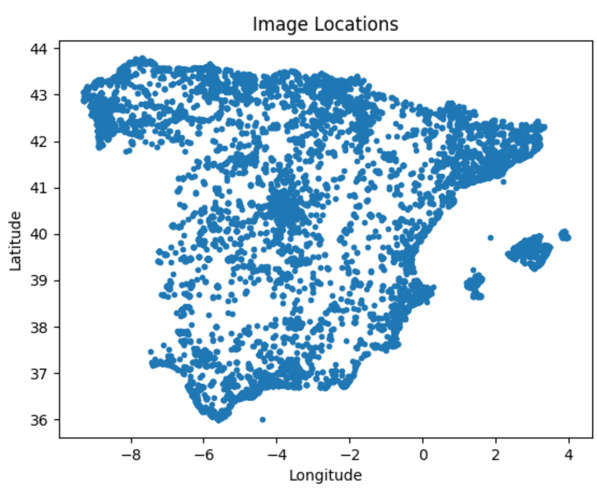

Figure 1: Geographic locations of images

Since the dataset consists solely of images within Spain, the plotted points resemble the map of Spain. Next, PCA was applied to the image features and the first two principal components were plotted, coloring each data point according to its longitude coordinate.

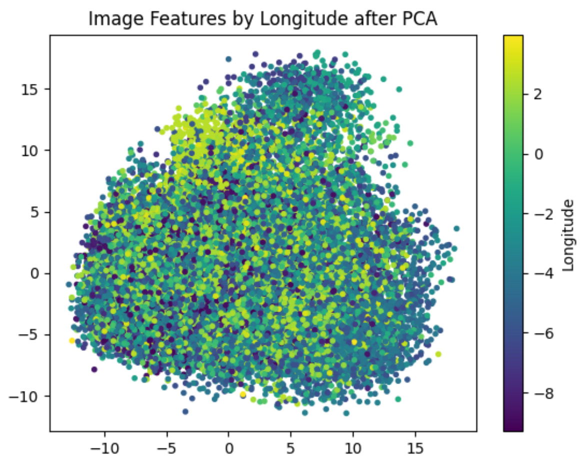

Figure 2: First two PCA dimensions colored by longitude

3 3-Nearest Neighbors

First, $k = 3$ was chosen to test the model. After training the 3-NN model on the training data, it was tested with an image of Palacio de San Telmo located at 37.380852N and 5.993192W.

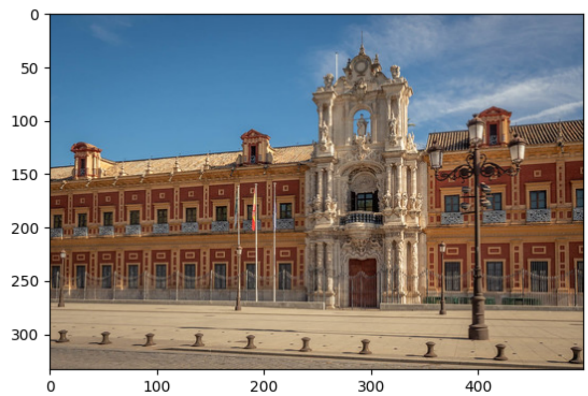

Figure 3: Image of Palacio de San Telmo

Then, the three nearest neighbors identified by the model were displayed.

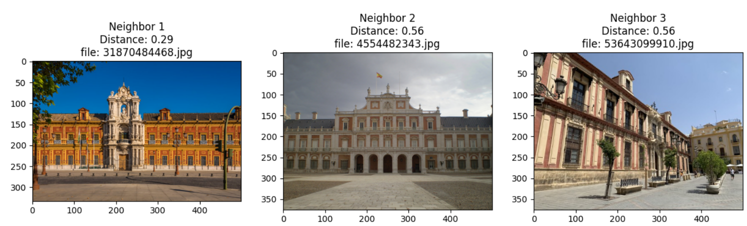

Figure 4: Three nearest neighbors of Palacio de San Telmo

The first image from Figure 4 is correct, with a distance of 0.29, while other images also seem to share similar features. Additionally, the file name, 31870484468, is the latitude concatenated with the longitude of the image, which closely matches the geographic location of Palacio de San Telmo at 37.380852N and 5.993192W.

4 Naive Baseline

A naive constant baseline was established by predicting the training set using the centroid. The mean displacement error (MDE) was 209.86266 miles.

5 MDE in k-NN

Next, a 1-D grid search with k-NN regression was conducted, varying k to find the optimal value. The MDE in miles was plotted to identify the best k.

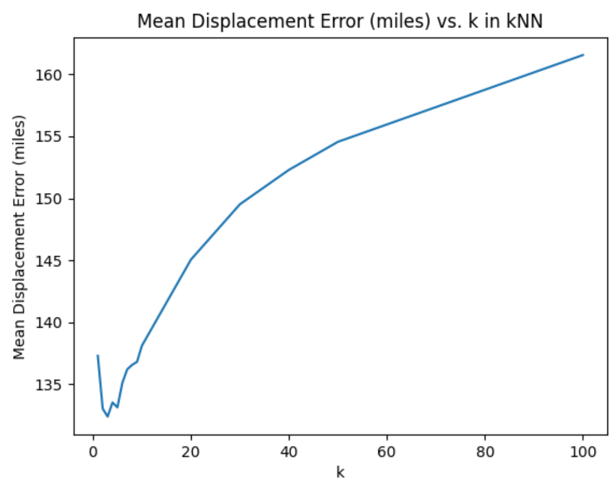

Figure 5: MDE with varying k neighbors

The graph shows that $k = 3$ yields the lowest error of 132.4 miles, which is an improvement from the model using the centroid.

6 Bias and Variance

By integrating bias-squared and variance over the probability distribution of photographs of Spain and given n training points, the results for k-NN models with different k were examined.

With $k = 1$, bias is low because the prediction is based on the single nearest neighbor. However, the model lacks generalization. As k increases, bias increases, and when $k = n$, bias is high as the prediction becomes the average of all training points, potentially missing relevant relationships between features and target outputs.

With $k = 1$, variance is high due to the model's sensitivity to the nearest neighbor, making it prone to overfitting. As k increases, variance decreases. When $k = n$, variance is low as the model becomes more stable by canceling out noise.

7 Weighted Neighbors

In the model, neighbors can be weighted differently, with closer neighbors receiving larger weights. Each neighbor was weighted by the inverse of its distance in feature space, and 10^{-8} was added to the denominator to avoid division by zero. The MDE for different k values was then plotted.

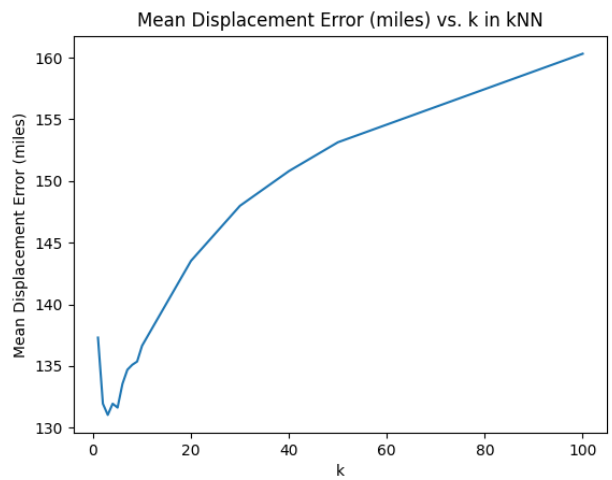

Figure 6: MDE with varying weighted k neighbors

Although the error did not drastically change compared to the naive k-NN model, there were improvements across all k values. The optimal $k = 3$ resulted in the lowest error of 131.9 miles.

8 Comparison between Non-parametric and Parametric Models

k-NN is a non-parametric model with complexity that grows with the amount of training data. In contrast, linear regression is a parametric model with a fixed number of parameters. The performance of k-NN and linear regression was compared at different training set sizes, using $k = 3$ for consistency.

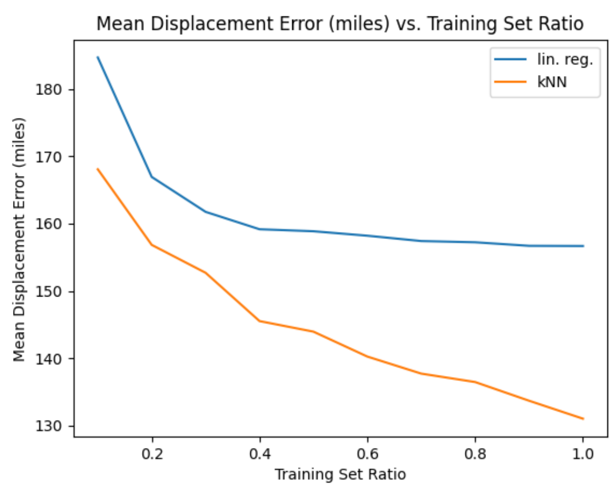

Figure 7: MDE comparison between k-NN and linear regression models

The graph shows that k-NN generally outperforms linear regression at all training sizes. The MDE for linear regression stabilizes at around 160 miles with 40% or more of the training data, whereas k-NN continues to improve with more data. Thus, with more training data, k-NN is expected to perform better than linear regression.

References

[1] James Hays and Alexei A. Efros. "im2gps: estimating geographic information from a single image". In: *Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2008.