ST ANDREWS SUPPORTING DOCUMENT

DANIEL PINIQI

Introduction:

Hello, my name is Daniel Piniqi and this is my document to support my application for direct entry to the Computer Science BSc at St Andrews University. I have attached my most advanced project in the language I am most experienced in – Python. I have attempted to thoroughly explain everything I have coded and I hope you enjoy reading through!

House Price Cost Prediction Programme:

Languages used:

* Written purely using Python

Purpose of Code:

* The purpose of this code is to approximate the cost of a house in London. I used a dataset for London housing with over 300,000 entries. This dataset was very unclean and I faced many errors dealing with missing values and unnecessary features. I ended up basing the prediction off of 4 main features: Square Footage, Living Rooms, Bathrooms and Bedrooms. Once I had played around enough to make the model equally efficient and accurate, I built a UI in TKinter.

Scale of Project:

* Approximately 200 lines of code
* 3 functions
* 0 classes

Use of Algorithms:

* Initially, I used linear regression as the model, however I found it led to extremely inaccurate values which I suspect is due to the lack of variety in my dataset and it’s inability to extrapolate. I then decided to settle for the Ridge regression model which led me to more accurate results and a lower margin of error. The Ridge regression model extends linear regression, helping to prevent overfitting by balancing the fit of the model with its complexity. I used the pandas data frame to ensure that my dataset was clean enough to pass through my model. The data set was extremely excessive and I found that initially, I was using too many features for the model that I ended up not needing. I used the simpleImputer function to fill in all the missing data by filling it up with the column’s median.

Conclusion:

I hope you enjoyed my explanation of the thought process that went behind this project and I look forwards to hearing from you soon! Thanks for reading.