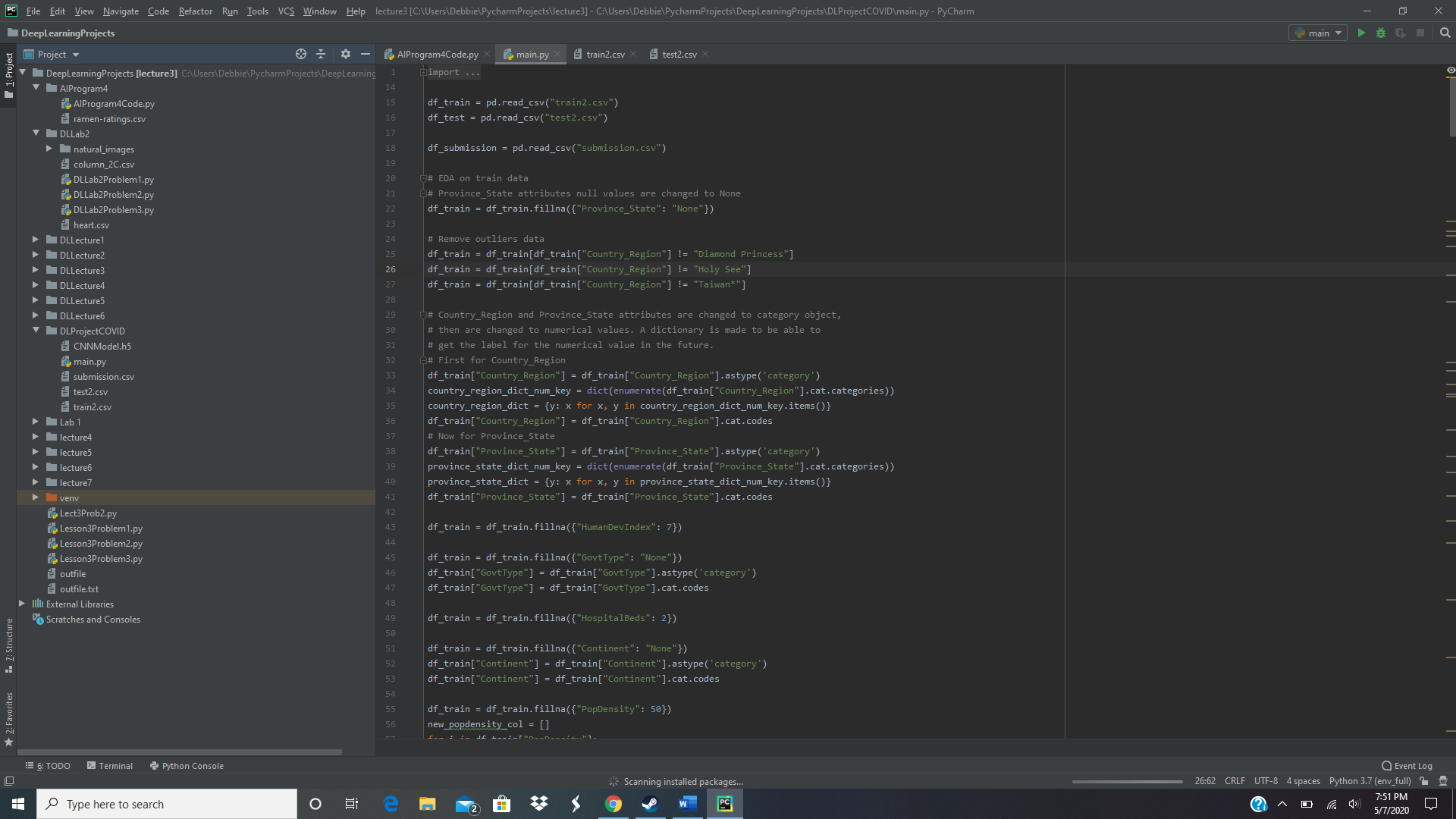
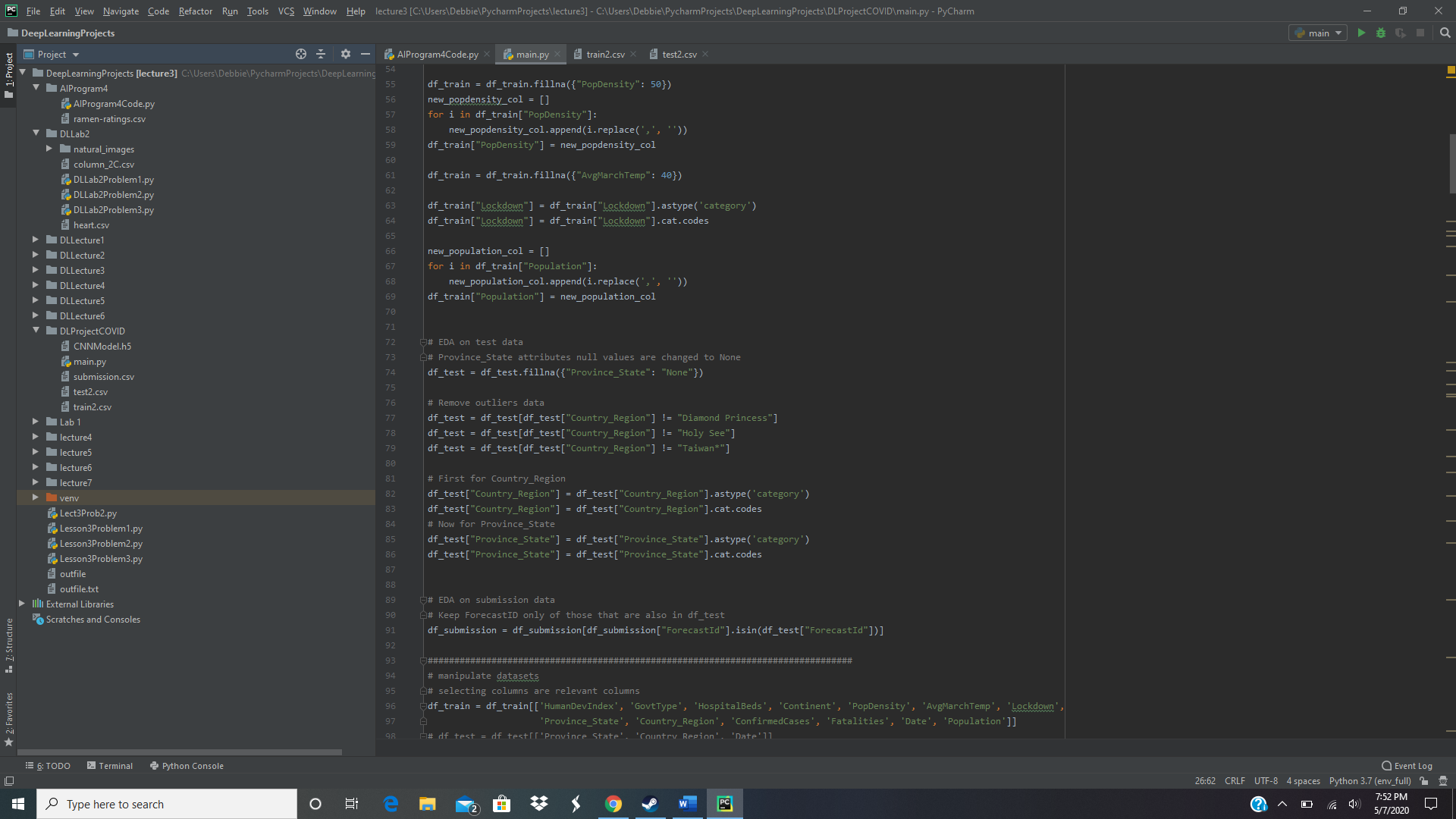
Deep Learning Final Project Increment 3

The dataset now consists of the 13 attributes from the previous increment, with an additional attribute added: total population. The final attributes are as follows:

* Province\_State – The province or state within a country or region. Many of these values are null.
* Country\_Region – Country or region.
* Date – Date in the form of yyyy-mm-dd.
* ConfirmedCases – Integer value of confirmed cases at the date in the location of the row.
* Fatalities – Integer value of fatalities at the date in the location of the row.
* HumanDevIndex – Human development index (indicates how developed a country is).
* GovtType – The type of government in the specified country.
* HospitalBeds – The number of hospital beds per 1000 people.
* PopDensity – The number of people per kilometer in the country or region.
* AvgMarchTemp – Average temperature in March in the specified country (Fahrenheit).
* Lockdown – A yes or no value, indicates whether the country or region went under lockdown.
* Population – The total number of people in the country or region.

An additional two columns were added in order to scale the confirmed cases and fatalities to the population of the region.

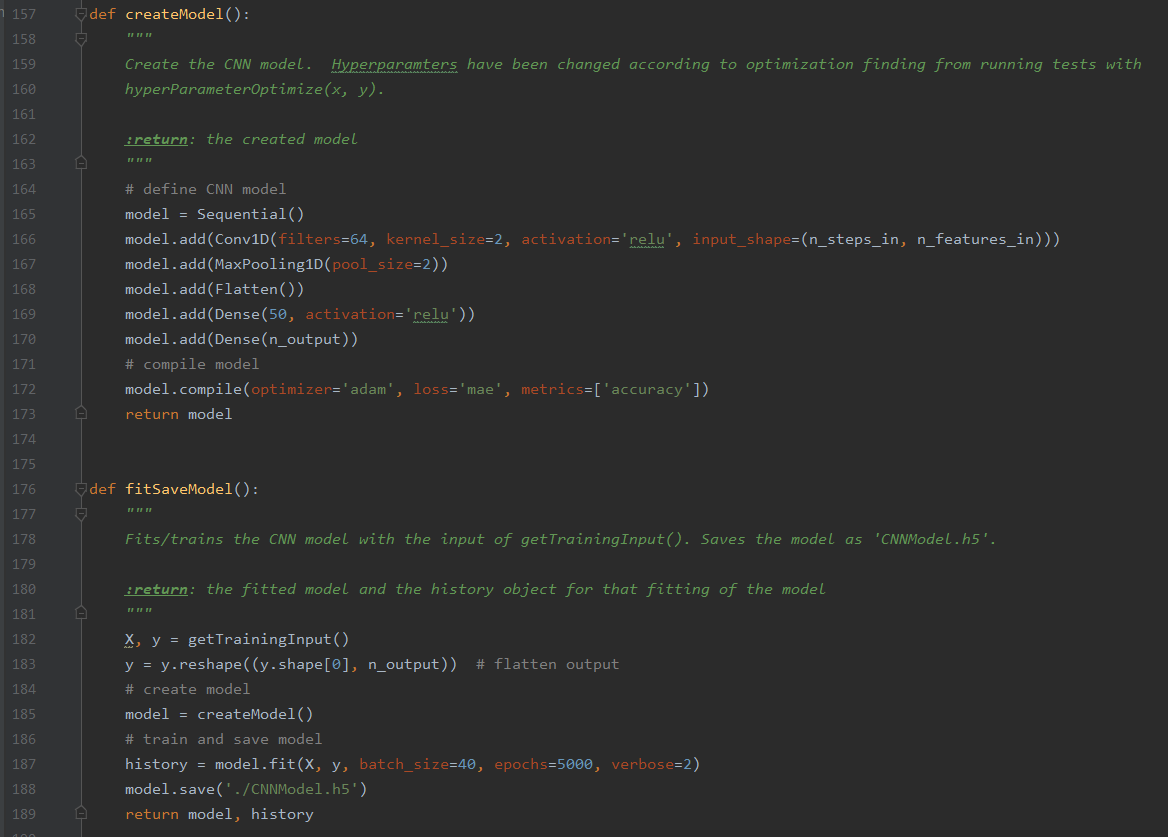




The current code for preprocessing the aforementioned data. Three small regions were removed due to insufficient data.

The model, which was an LSTM in increment 2, was changed to a CNN. The CNN was able to handle non-sequential data in addition to the main sequential data, and both types of data are now used in the network.

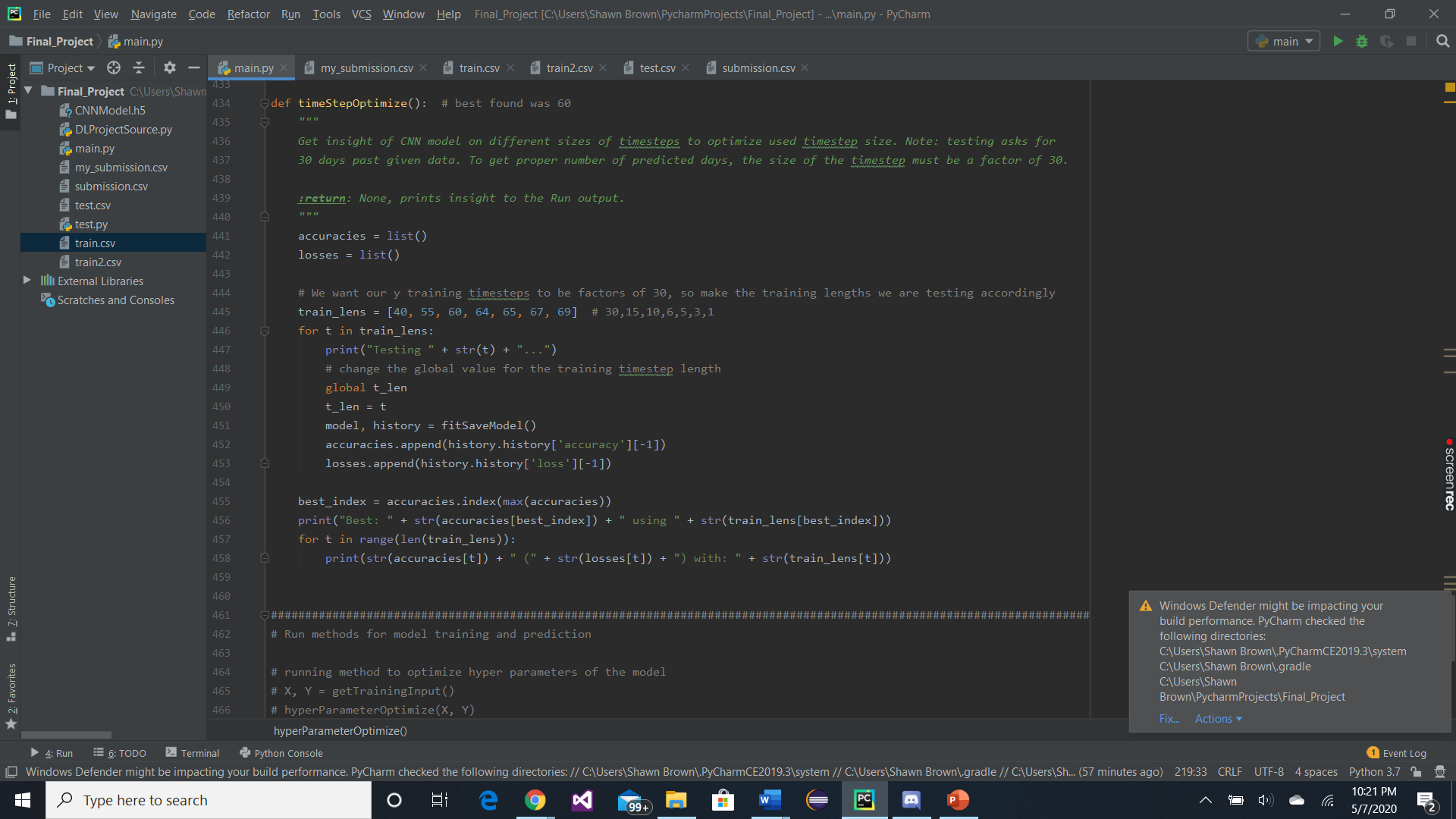
CNN Model:



Model predicts in timesteps of 10, meaning 10 future days of data are predict from model.predict().

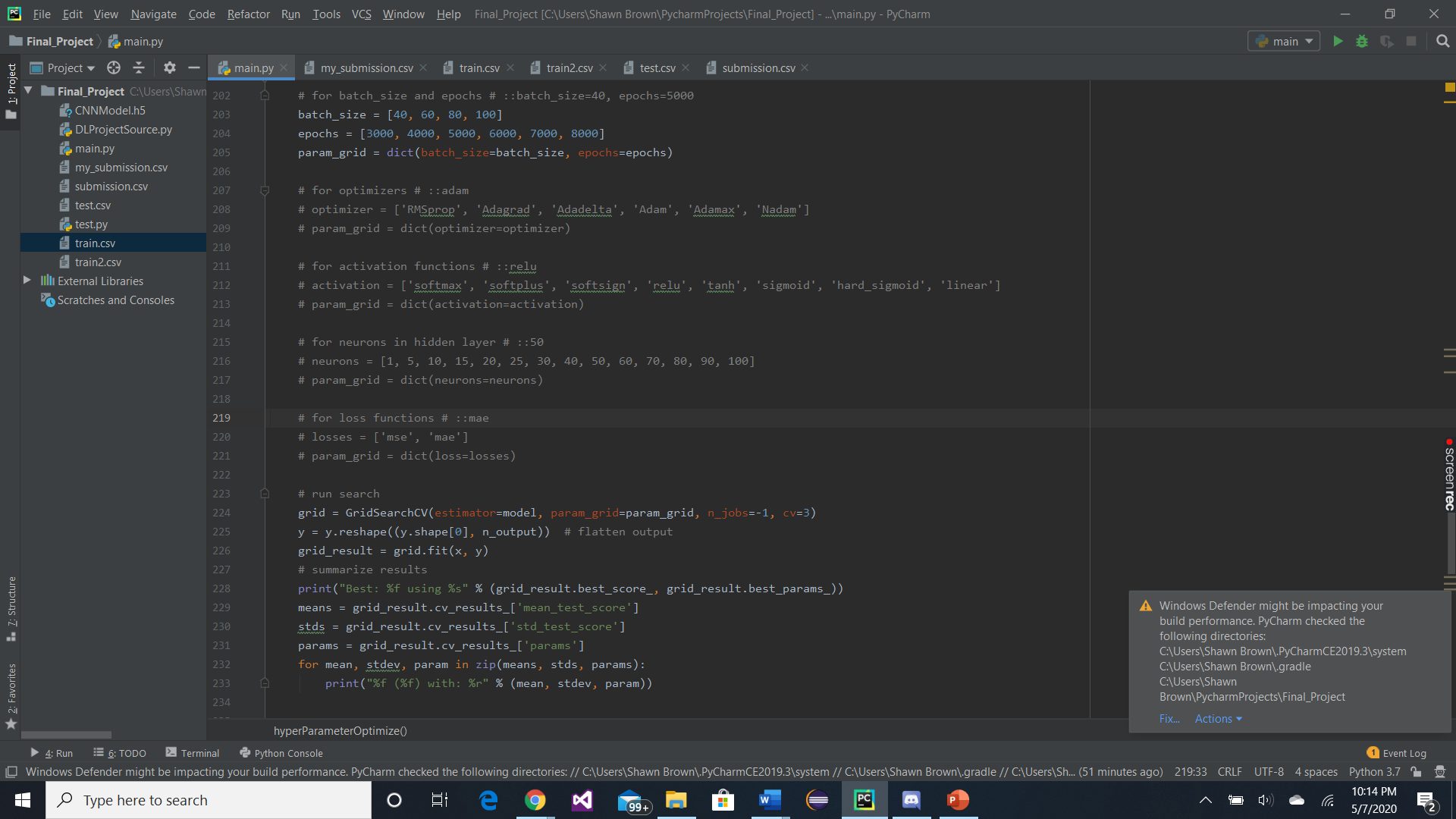
Timesteps options (30,15,10,6,3,1) were search and optimized by creating method timeStepOptimize().

Code:



Optimized hyperparameters of the model using GridSearchCV.

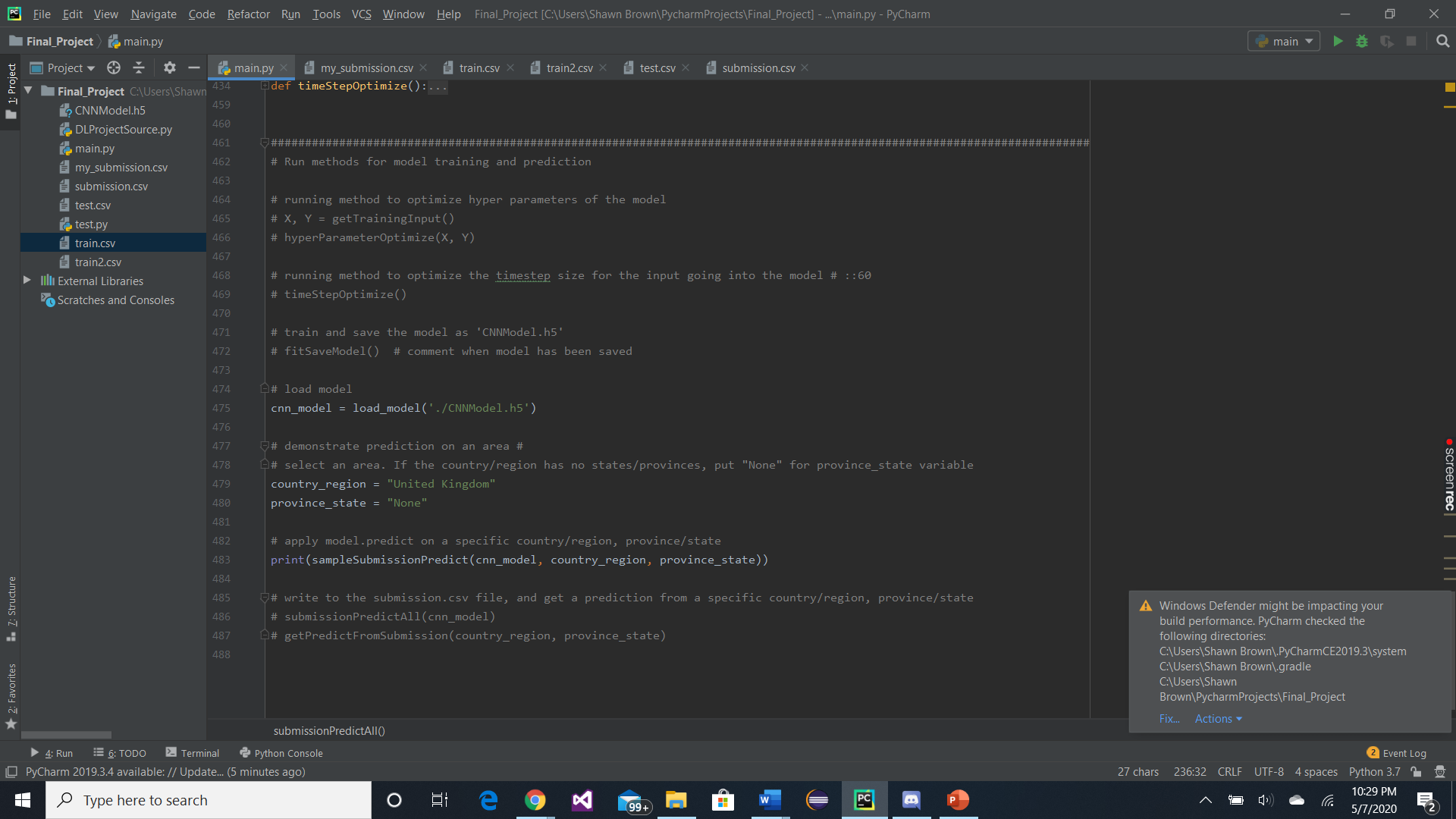
Code:



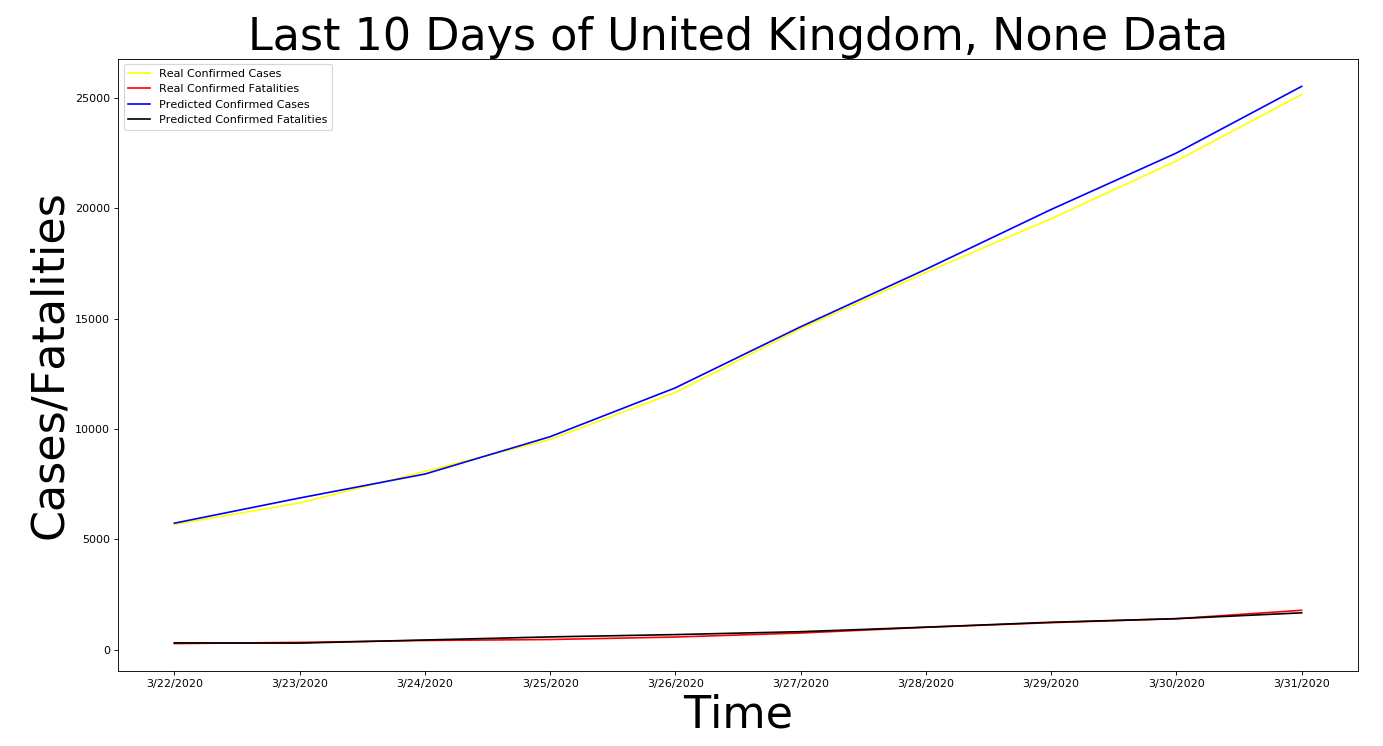
Optimized:

* Batch\_size: 40
* Epochs: 5000
* Optimizer: adam
* Activation: relu
* Neurons in hidden layer: 50
* Loss function: mae

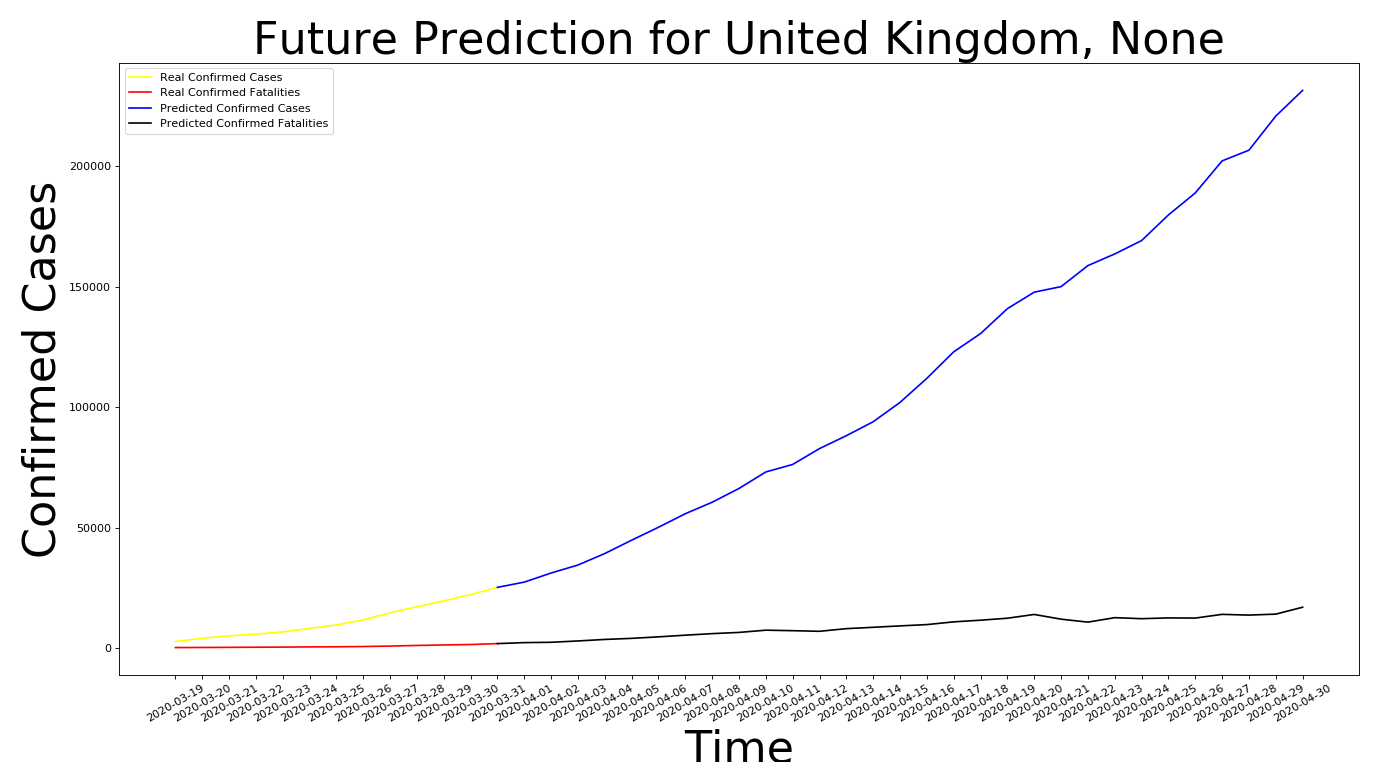
Can have to model predict and plot data for a specific area:



Show last 10 days of known data prediction for comparison to actual data:



And the future data predicts:



Able to predict all data needed to go into the submission.csv file. The submission file is written to with this data using method submissionPredictAll(model). Can get data found of specific locations with getPredictFromSubmission(country, province).

