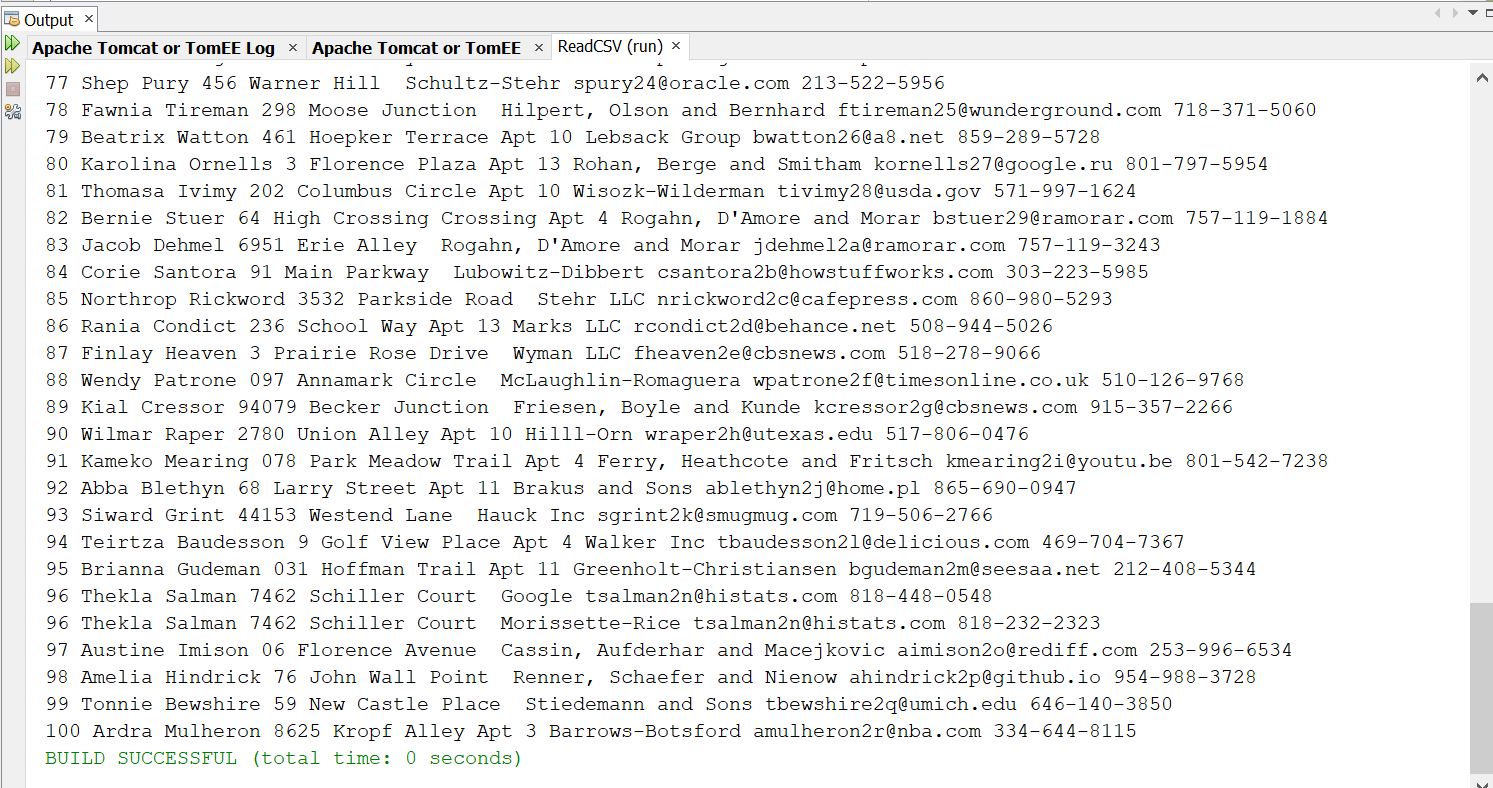
Working ScreenShots:

Part (A)

---CSV read data



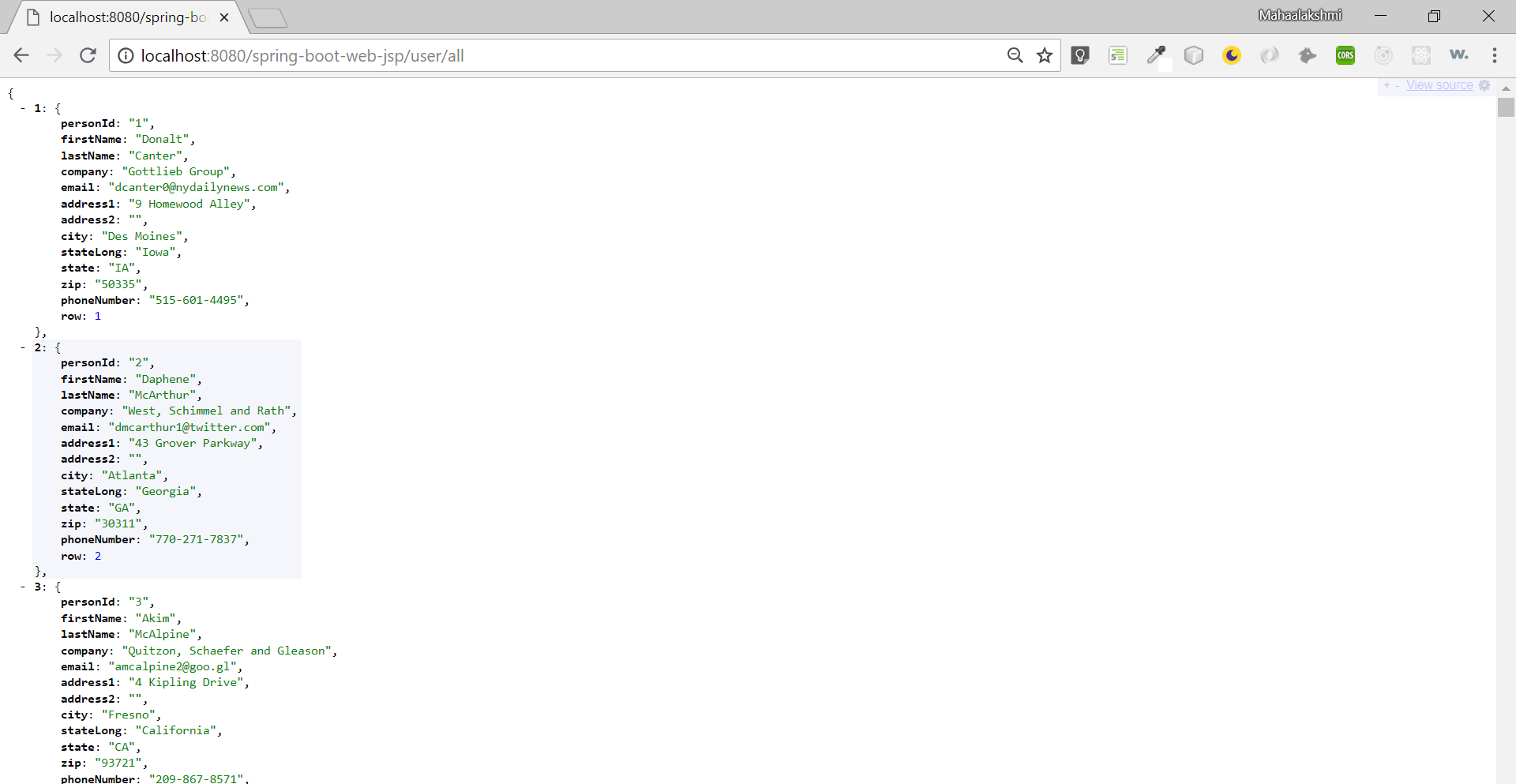
Part (B):

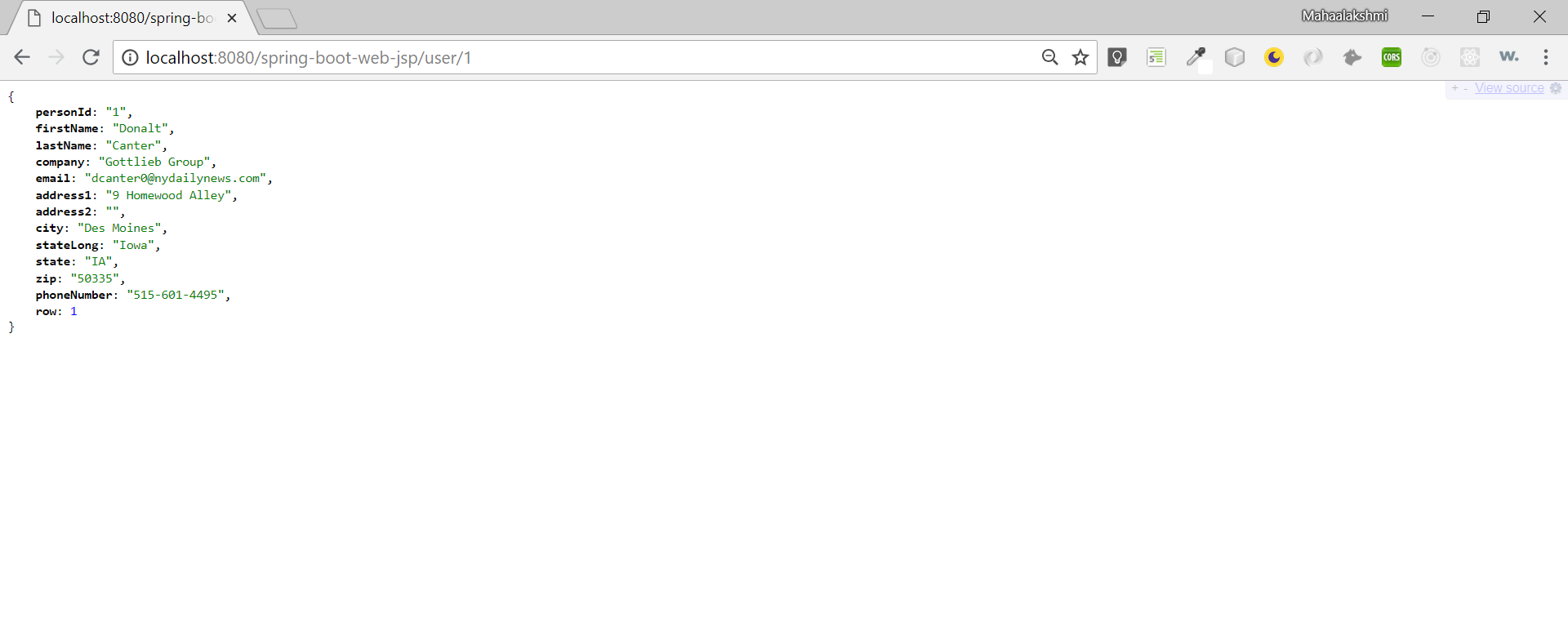
Brief Overview –

The web application is built in Spring Boot, it has two major functionalities to identify duplicate records from CSV file and to classify using trained model weights as duplicates or non-duplicates.

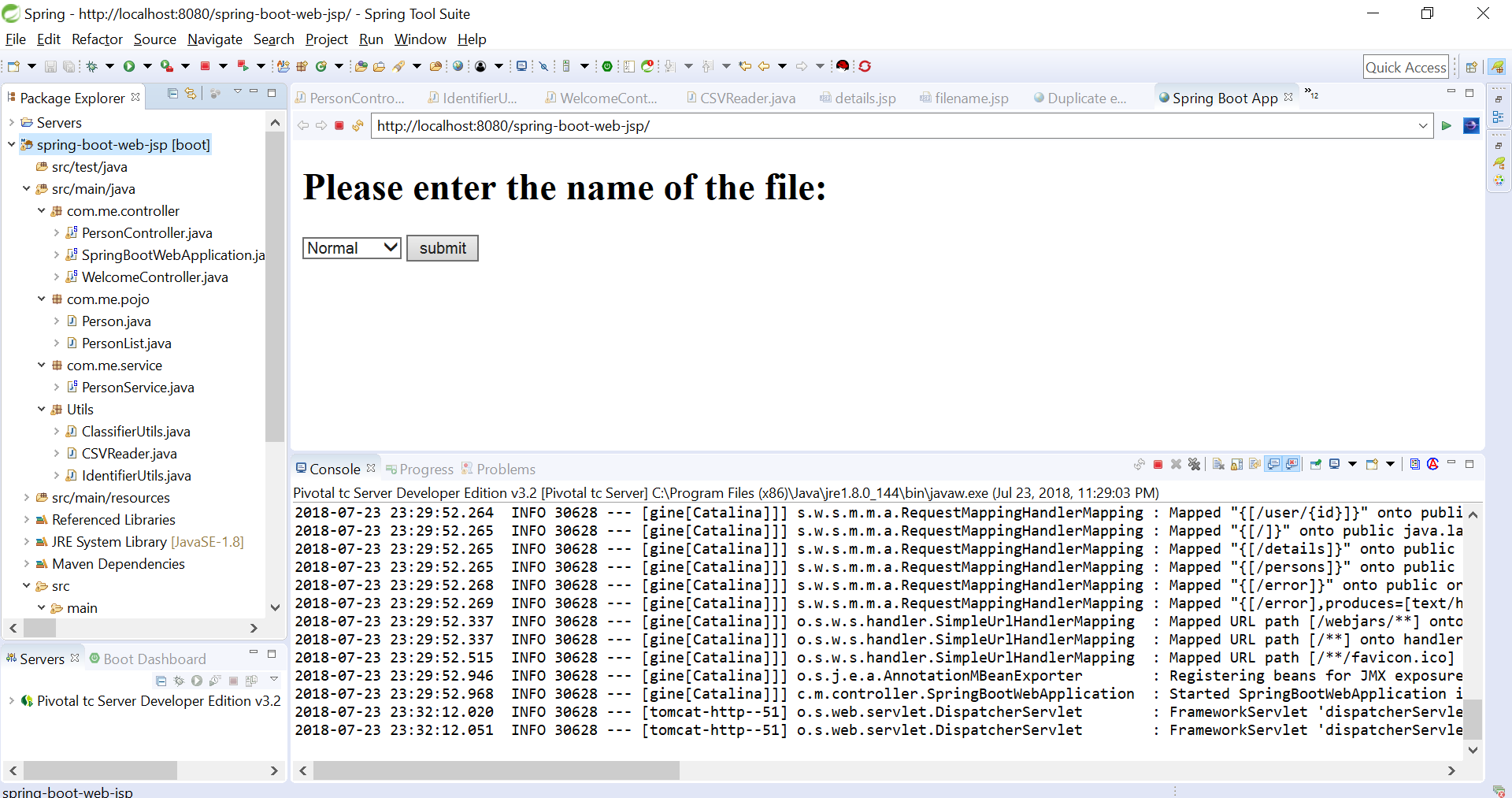
I have built a single layer perceptron in this project, to calculate build the model, using weights calculated using metaphone (phonetic similarity identifier library) output of the Company Names, and then transferred the model learning to identify, duplicates using emails and calculated the expected and observed weights which are similar but not yet accurately.

JSON Object Accesses:

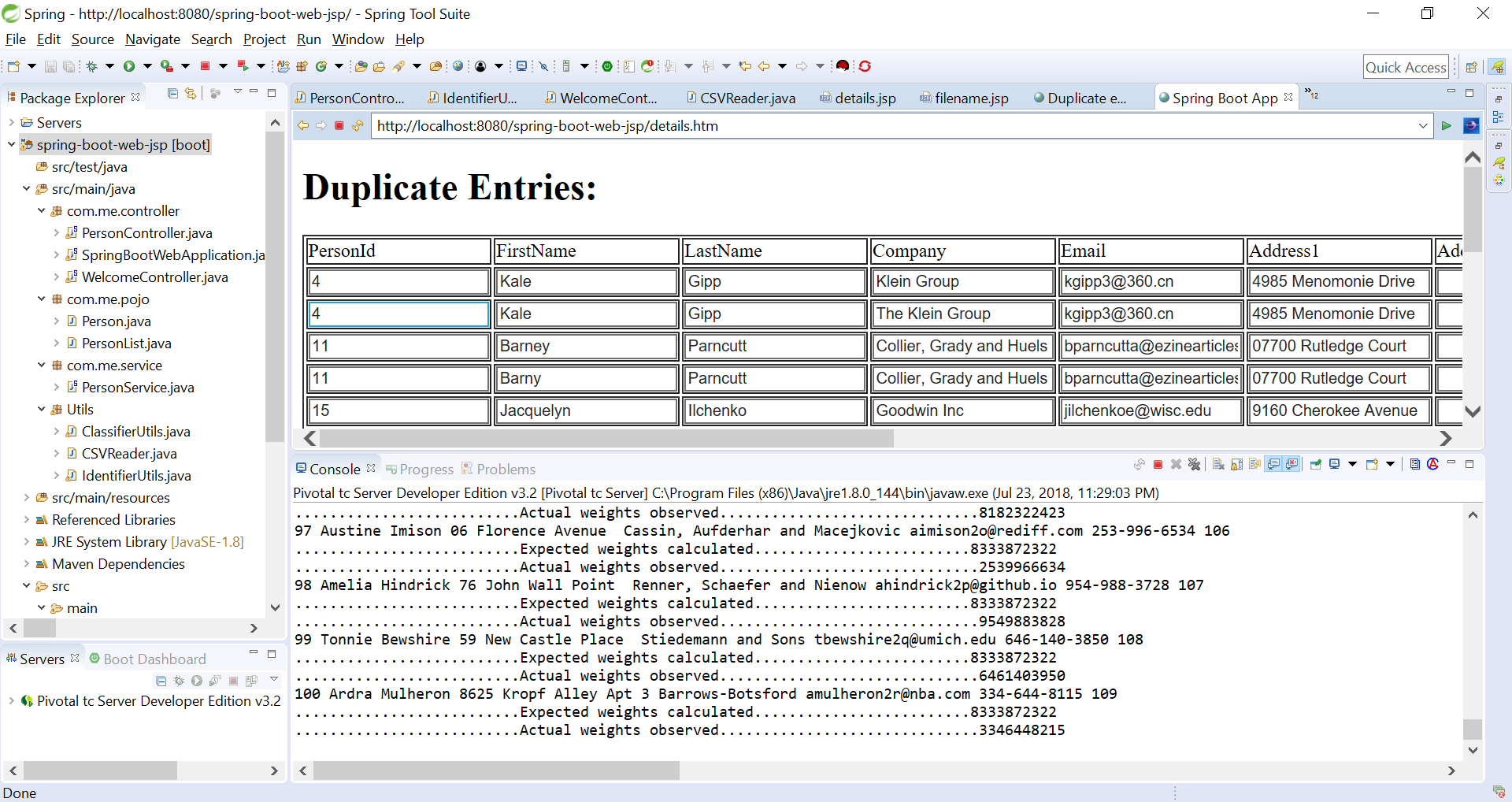




Screen 1 -



Screen 2 -



To scale this up further, mapReduce could be used.