**CS539 Machine Learning. Fall 2014. Project 3: Neural Nets. Students’ Names: <write your names here>**

**[8 points] Dataset 1 Description and initial data preprocessing if any (at most ½ page):**

For the S&P 100 Companies, the daily adjusted closing prices of every company were downloaded as .CSV files using an automated Quote Downloader Excel notebook from Invest Excel (http://investexcel.net/). Because each company had its own .CSV file, these files were combined to one .CSV file using a custom Perl script. This script also converted the dates to days since Jan1, 2013, and put the **BIIB** prices last, so they would be automatically detected as the target attribute in WEKA. This process was repeated to acquire the training data. There was no missing data except for the **GOOG** prices before the Google stock split on March 27, 2014 (http://investorplace.com/2014/07/goog-google-stock-split/), so **GOOG** was removed from both the training and test data.

**[4 points] Three Guiding Questions (at most ½ page):**

1. How does **BIIB** stock fluctuate with changes in **ABT** (Abbott Laboratories, a large pharmaceuticals company)?
2. How does **BIIB** stock fluctuate with changes in **AMGN** (Amgen Corp., another large biotech company)?
3. How does knowledge of **time** (Date or days since Jan 1 2013) affect the accuracy of prediction.

**[8 points] Dataset 2 Description and initial data preprocessing if any (at most ½ page):**

**[4 points] Three Guiding Questions (at most ½ page):**

**1. …**

**2. …**

**3. …**

**Code (at most 1 page)**

**[10 points] Description of Weka Code (at most 1/2 page):**

**[15 points] Description of R Code, Packages, and Functions (at most 1/2 page):**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Dataset 1. Weka Experiments.** *at most 1/2 page.* | | | | | | | | |
| **Guiding**  **questions** | **Train/**  **Test** | **Pre-process** | **Net Topology** | **Parameters** | **Confusion matrix**  **& Accuracy** | **Time to build model** | **Analysis & observations about experiment** | **You can add other columns** |
| 1? 2? 3? | 10-fold cross-val? |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Dataset 1. R Experiments.** *at most 1 page.* | | | | | | | | |
| **Guiding**  **questions** | **Train/**  **Test** | **Pre-process** | **Net Topology** | **Parameters** | **Confusion matrix**  **& Accuracy** | **Time to build model** | **Analysis & observations about experiment** | **You can add other columns** |
| 1? 2? 3? | 10-fold cross-val? |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

**Summary of Results and Visualizations for Dataset 1 (at most 1/2 page)**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Dataset 2. Weka Experiments.** *at most 1/2 page.* | | | | | | | | |
| **Guiding**  **questions** | **Train/**  **Test** | **Pre-process** | **Net Topology** | **Parameters** | **Confusion matrix**  **& Accuracy** | **Time to build model** | **Analysis & observations about experiment** | **You can add other columns** |
| 1? 2? 3? |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Dataset 2. R Experiments.** *at most 1 page.* | | | | | | | | |
| **Guiding**  **questions** | **Train/**  **Test** | **Pre-process** | **Net Topology** | **Parameters** | **Confusion matrix**  **& Accuracy** | **Time to build model** | **Analysis & observations about experiment** | **You can add other columns** |
| 1? 2? 3? |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

**Summary of Results and Visualizations for Dataset 2 (at most 1/2 page)**

**Overall Conclusions, Observations, and Visualizations for Project 3 (at most 1/2 page)**

**Authorship: Describe in detail the work that each of the two team members did on this project. This authorship description must accurately reflect the work done by each team member, and must be approved by the two members of the team (at most 1/2 page)**