# Length of Stay Prediction

# Weekly Project Meeting Minutes

*The main purpose of the document is to capture all the work that has been done by the group over the course of one week and* ***not*** *to write down what was discussed in a single meeting. You should be meeting and/or working throughout the week*.

**Time group spent on project: 6 hours**

**Group Number: 11**

**Group members present (Name, ID):**

* **Varinderjit Singh (0730482)**
* **Kanchan Bagga (0732356)**
* **Jaspreet Kaur (0730470)**

**Specific Activities from prior week:**

* **List brief description of activities carried out by group member**

**Varinderjit Singh:** Python coding which he did the modelling and investigate that data in SQL.

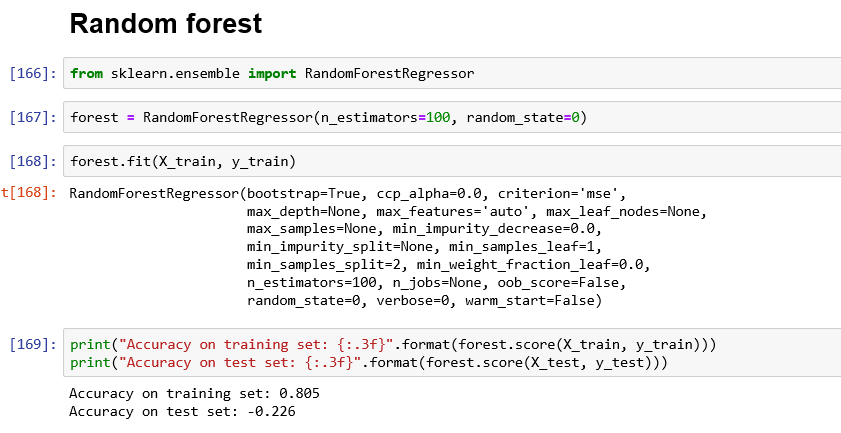
**Kanchan Bagga:** She is working on project report. Reading the different article related neural networks on our project which will be helpful in future.

**Jaspreet Kaur:** She is reading the article on neural networks and some coding which related the neural network. Yet coding part is not proper work, but she will be tried.

**Specific Output from prior week:**

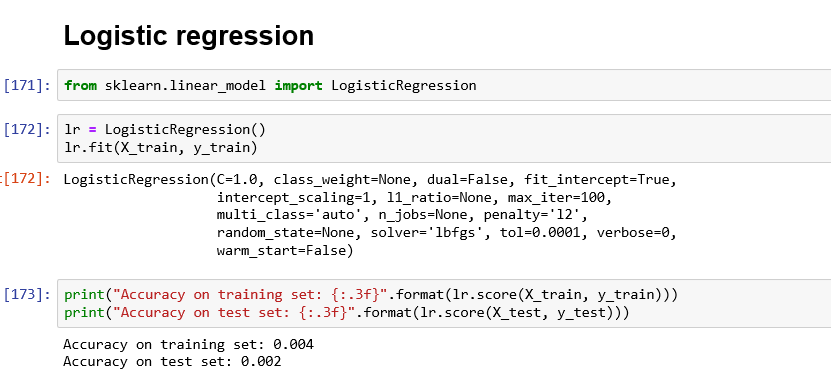
* **Include brief summary of any written work, experiments, or code developed**
* **Attach actual output as a separate file when submitting minutes; for example, export your Jupyter notebook as an html file and upload that with your minutes**

Firstly, we do the random forest model using the regression technique which we get the training accuracy 0.805 and the test accuracy -0.226 it is very less. So, we tried other models on our dataset.

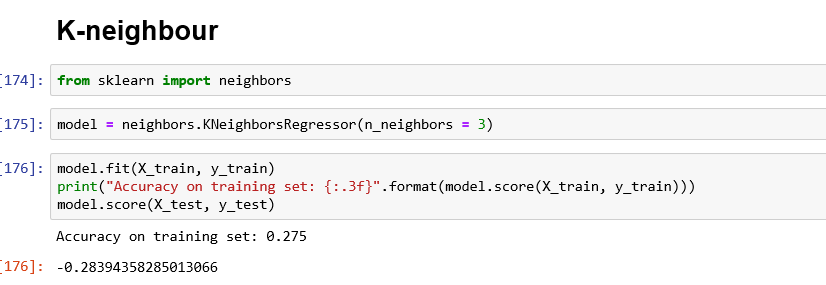


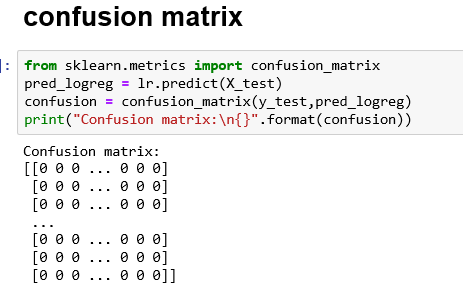
This is our second model which gradient boosting to train the accuracy that is 0.482 on training and on test is 0.026..



We also tried logistic regression to get the better accuracy from last two see the accuracy which is training set 0.004 and testing accuracy is 0.002

When we implement the K- neighbor model as regressor the accuracy of training set is 0.275 and the test set is -0.283.





|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Project Contribution | | | | | |
| **Major Task** | **Member Name** | **Sub Tasks** | **Description** | **Completed on** | **Status** |
| **Topic selection** | Kanchan Bagga | Finding project topic, Topic related search | Project that could be useful in future. Should have something new. Scope for job placement. Should be useful in real life as a tool. | February 04, 2020 | Done |
| **Dataset** | Jaspreet Kaur | Finding dataset (Searching, getting access and downloading) | Searched dataset from different websites. Sent emails to many of the organizations for accessing their datasets. Got access for MIMIC dataset and downloaded | February 18, 2020 | Done |
| Jaspreet Kaur  Kanchan Bagga  Varinderjit Singh | Target and supporting variables selection | Target variable (LOS) was already cleared in everyone’s mind but all the supporting variables were needed to select that could effect on LOS. | February 25, 2020 | Done |
| Kanchan Bagga | Data Cleaning (Changing format, keeping necessary data only) using excel | Changed downloaded files format. Made one single file from 27 files by merging all necessary columns in master file only. Removed unnecessary columns only, removed outliers, data about emergency department only was kept. Some new columns are made. Missing data rows are removed. | March 03, 2020 | Done |
| Varinderjit Singh | Cleaned data understanding & exploration | Understood & explored cleaned data in SQL Server for my satisfaction and cross checking. | March 05, 2020 | Done |
| **Project related research** | Jaspreet Kaur | Searching work done on related projects | Projects that are already made on same topic. So that we could come to know what is done and what different we can do. | March 03,2020 | Done |
| **Article reading** | Varinderjit Singh  Kanchan Bagga  Jaspreet Kaur | Studied 10 (4-3-3) articles individually | Study articles based on how we can use machine learning in our project. For understanding what kind of models, we can use and how we can improve their accuracy. | March 12, 2020 | Done |
| **Model Study** | Jaspreet Kaur | ANN, Logistic regression,  K-neighbor | YouTube tutorials,  Google search | March 23, 2020 | Done |
| Kanchan Bagga | CNN, Random forest, Gradient boosting | YouTube tutorials,  Google search | March 19, 2020 | Done |
| Varinderjit Singh | Confusion matrix, Multiple linear regression | YouTube tutorials,  Google search | March 20, 2020 | Done |
| **Model Building** | Varinderjit Singh | Neural network,  Multiple linear regression,  Confusion matrix |  | - - - | Pending |
| Kanchan Bagga | Random forest,  Gradient boosting |  | March 24, 2020 | Done |
| Jaspreet Kaur | Logistic regression,  K-neighbor |  | March 24, 2020 | Done |
| **Report writing** | Jaspreet Kaur |  |  | - - - | Pending |
| **Presentation** | Kanchan Bagga |  |  | - - - | Pending |

**On Target:**

* **Indicate the current status of your project**
  + \_\_\_\_\_ green: everything on track for completion by due date

**Challenges/Disagreements:**

* **List any particular challenges identified/discussed and possible solutions**
  + **include tasks causing a yellow or red flag for your project**
* The main challenges of us neural network. It is difficult to implement We are tried but not good result. We implement the different models technique but we did not get good accuracy because most of accuracy was negative how our model get the better accuracy which is main problem when we do the these models.
* List any notable disagreements and subsequent discussion and resolution

**Planned Activities for coming week:**

* **List brief description of activities by group member**
* **Make sure tasks are assigned to address yellow and red flag items**

Neural networks models will be tried to implement (Varinder)

(Jaspreet and Kanchan) also perform the work on reading the article on neural networks which is more helpful to implement the modeling on our project.

Also, we will back to back report writing.