**Question3: L**ikelihood of “major injury” from a fall based on available factors was determined through first cleansing the data in excel and taking only relevant predictors into the python’s Anaconda so that data modeling algorithm can be applied for predicting the factors responsible for Major Injury.

Out of variables mentioned , some were kept and some were neglected as per the reasons mentioned next to them.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Fall\_SKey | Witness\_Flag\_key | Cohorent\_Flag\_Key | InJuryLevel\_Key | Major Injury Key | Location\_Key | PatientActivity\_key | Latest WHeFRA score | Functional\_Key | Medication\_Key1 | Medication\_Key2 | Medication\_Key3 | Medication\_Key4 | Medication\_Key5 | medication\_impact | History\_Fall | NurseActivity\_Key | PatientUnitCase\_key | Nurse\_Key | Date\_key | dateName | Weekday\_Key | Time\_key | Fall | Shift\_key | overlap | WHeFRAscore\_key | MedicationNos | NurseHourseonShift | TimeofFall |

|  |  |  |
| --- | --- | --- |
| Variable | Used for final Modeling | Reason for not taking up for final modeling |
| Fall\_SKey | No | It is just a serial number |
| Witness\_Flag\_key | Yes |  |
| Cohorent\_Flag\_Key | Yes |  |
| InJuryLevel\_Key | No | This key was converted into a dummy key which is Major Injury key as prediction was required for Major Injury |
| Major Injury Key | Yes, This was the dummy variable crated. This is the Y or Dependent Variable |  |
| Location\_Key | Yes |  |
| PatientActivity\_key | Yes |  |
| Latest WHeFRA score | No | The key for using this is there in another variable named “WHeFRAscore\_key” |
| Functional\_Key | Yes |  |
| Medication\_Key1 | No |  |
| Medication\_Key2 | No |  |
| Medication\_Key3 | No |  |
| Medication\_Key4 | No |  |
| Medication\_Key5 | No |  |
| medication\_impact | This is the new variable that was created by knowing the impact that each medication can cause to the patient | Logic used was> If even one medication was given that was sleep inducing , then medication impact was much more likely to cause injury to the patient. |
| History\_Fall | Yes |  |
| NurseActivity\_Key | Yes |  |
| PatientUnitCase\_key | No | It gives just a serial number to patients |
| Nurse\_Key | Yes |  |
| Date\_key | No | This key doesn’t give any data for prediction |
| Weekday\_Key | Yes |  |
| Time\_key | Yes |  |
| Fall | No | Because this says all of the cases were falls |
| Shift\_key | Yes |  |
| overlap | Yes |  |
| WHeFRAscore\_key | Yes |  |
| MedicationNos | Yes |  |
| NurseHourseonShift | Yes |  |
| TimeofFall | No | This is just a date which doesn’t add value to predict major injury probability |

In python, logistic regression was first applied but since it did not yield right results, then random forest classifier algorithm was applied which yielded very good results which are summarized under:

precision recall f1-score support

0 1.00 1.00 1.00 952

1 1.00 1.00 1.00 164

avg / total 1.00 1.00 1.00 1116

Hence precision of predicting the likelihood of having a major injury or no major injury was 1 for test data which was :

0 952

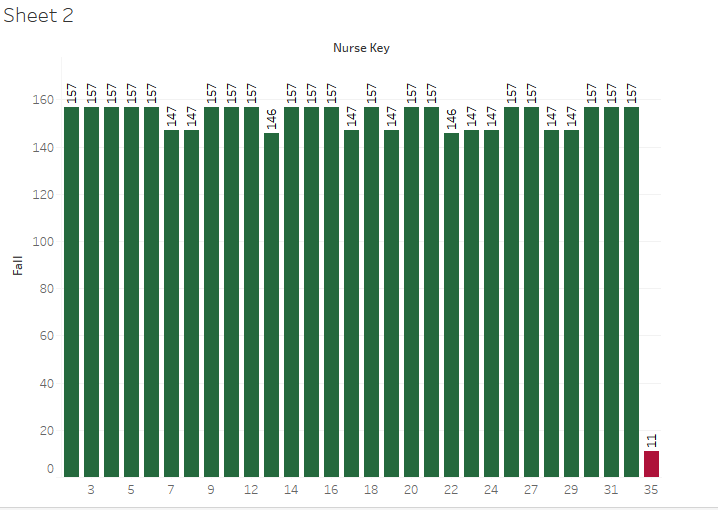
1 164

Then features/predictors were ranked according to the significance they had in causing a major injury:

|  |  |
| --- | --- |
| **Feature/Predictor** | **Importance** |
| **Nurse\_Key** | 0.164951 |
| **NurseActivity\_Key** | 0.100264 |
| **Time\_key** | 0.095833 |
| **Functional\_Key** | 0.094700 |
| **History\_Fall** | 0.090108 |
| **PatientActivity\_key** | 0.086677 |
| **Location\_Key** | 0.082905 |
| **NurseHourseonShift** | 0.076740 |
| **MedicationNos** | 0.041712 |
| **Witness\_Flag\_key** | 0.039396 |
| **WHeFRAscore\_key** | 0.035128 |
| **Cohorent\_Flag\_Key** | 0.033552 |
| **Shift\_key** | 0.027971 |
| **overlap** | 0.018128 |
| **medication\_impact** | 0.007197 |
| **Weekday\_Key** | 0.004738 |

Hence looking at the outcome, the fall prevention mechanism suggested is:

1. The nurses in the hospital need a training on handling patients of respective ailments. Almost all of the nurses had falls under their supervision. Only Nurse 35 has had least number of falls. Hence she should be used as a benchmark for training other nurses.

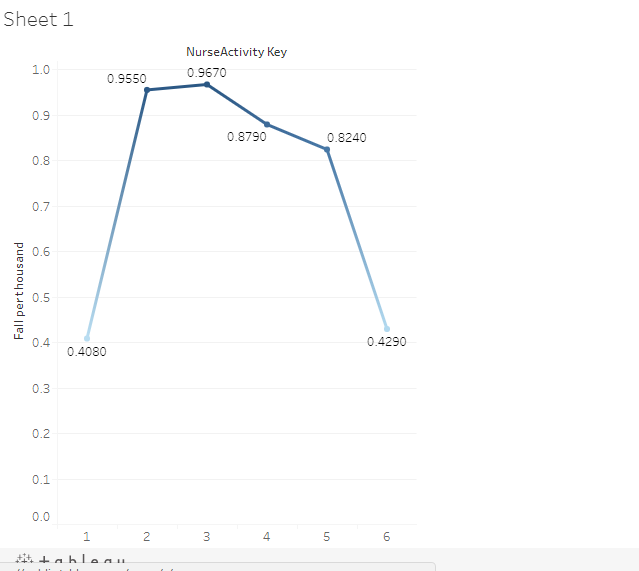


1. Nurse activity which caused major injuries were : 2,3,4 corresponding to

***was helping patient in stationary activity****>> This case needs serious analysis as this means nurse administering the patient during this activity had serious problems.*

***was with another patient****>> If the nurse was supposed to be attending the patient in case and still was with another patient, this means she is not performing her duty well.*

***was at the nurses station****>> This means she was not attending the patient when she was supposed to. This also seems to be lack of duty n part of the Nurse.*



1. Time of the day when most injuries happened was: 3 to 5 pm and 9 to 10 pm and 3 to 4 am as show in the below figure. Hence these time hours require strict supervision.

