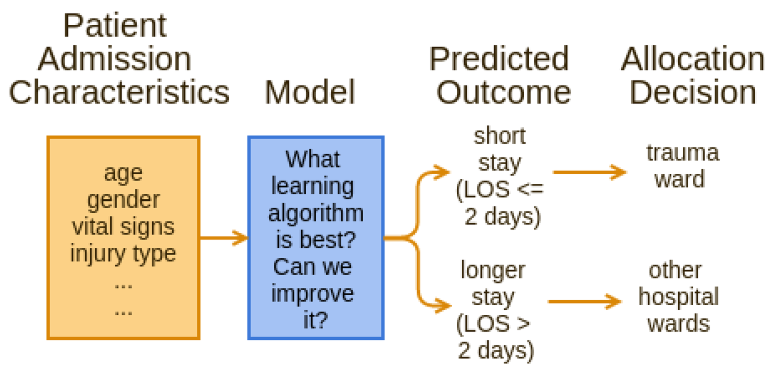
IMPROVED PREDICTION OF HOSPITAL LENGTH OF STAY FOR SEVERE INJURY

Tianyu Pu tianyu.pu@sydney.edu.au
Irena Koprinska
School of Information Technologies
Faculty of Engineering and Information Technologies

1. Problem

There are limited beds in hospital trauma wards, and yet there is a constant demand for these beds by the inflow of severely injured patients. Many patients are initially allocated to these beds when they could be better treated in another specialised ward.



If we could accurately classify patients with hospital length of stay (LOS) of 2 days or less versus those who require longer stays, we could make a more informed decision whether or not to place them in another ward when they are admitted, rather than wasting time and resources transferring them to another ward later.

2. Motivation

Accurate prediction of the LOS in various medical domains (such as burns) has been extensively studied as it is a key indicator of how hospitals use their limited resources. The current literature focuses extensively on generalised linear models, with some applications of techniques from machine learning such as neural networks. There has been no work done in systematically applying a range of machine learning techniques to LOS prediction, and specifically not to the trauma medical domain. Additionally, there has been little work done in feature transformation and selection when LOS prediction models are created.

3. Contribution

We systematically investigate feature transformation and selection techniques in the construction of a LOS prediction model for trauma patients. We also apply and evaluate a comprehensive range of classification algorithms on data from the trauma domain as well as from a general hospital setting.

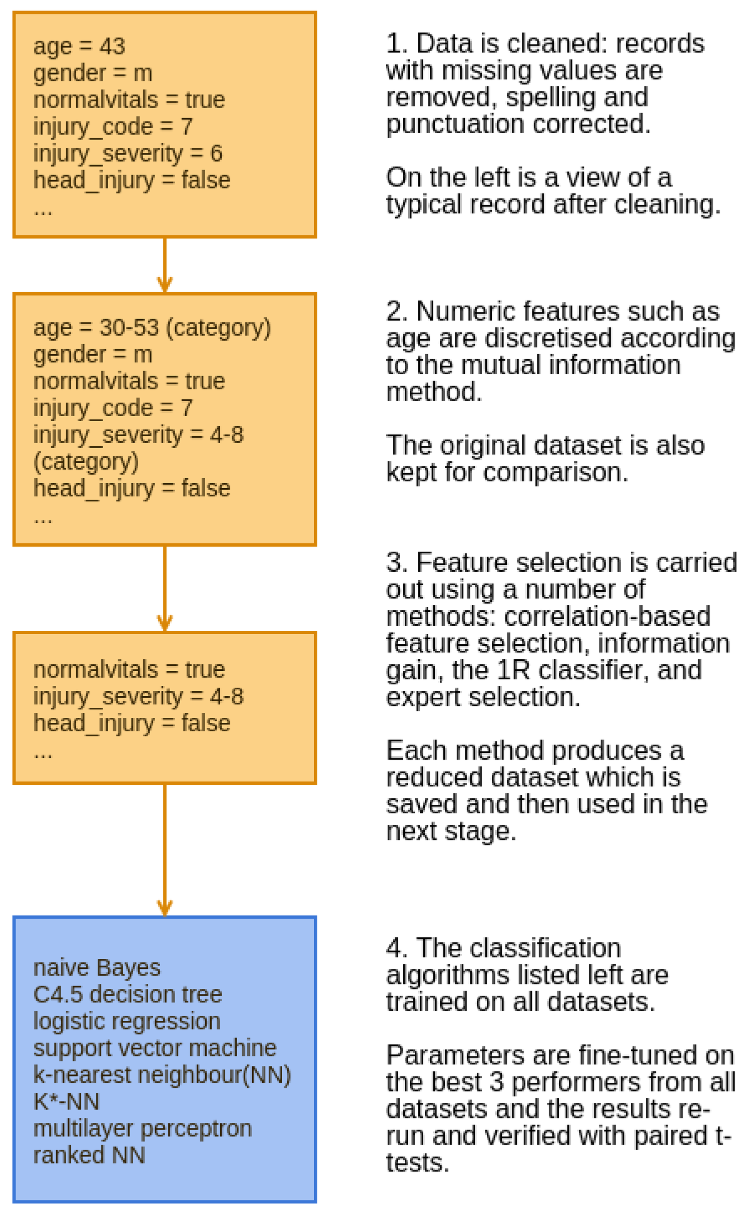
In addition, we propose a new nearest-neighbour (NN) algorithm, ranked NN, which takes into account the predictive relevance of features when computing the distance to the nearest neighbours.

4. Datasets

Our study was conducted on two datasets: one with 2546 records from the Trauma Services Centre at the Royal Prince Alfred Hospital in Sydney, consisting of trauma patients admitted to the centre between 2007–11; the other from the Hospital das Foras Armadas in Portugal with 17546 records collected from 2000–13 and covering a wide range of medical diagnoses.

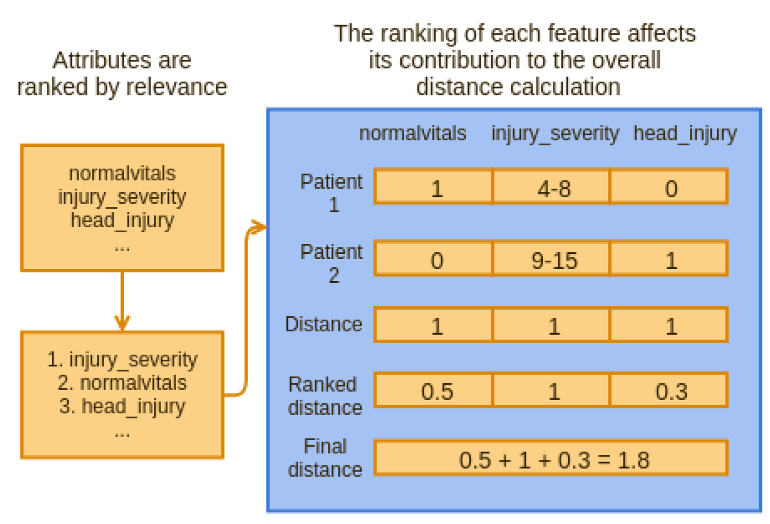
5. Approach

5.1. Feature preprocessing and investigation of learning algorithms



5.2. Ranked nearest neighbour

Idea: the contribution of a feature to the distance between two records should be proportional to its level of 'relevance' or predictive power with respect to the outcome variable, LOS. This is how it works:



6. Evaluation Metrics

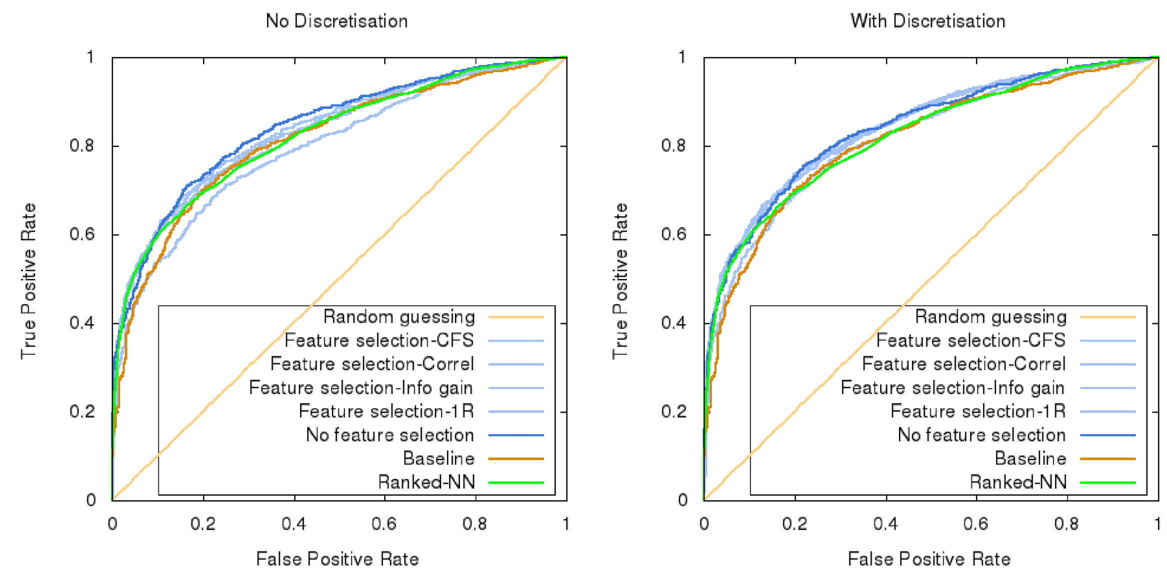
*Area under the receiver-operating characteristic curve (*AUC*)*, the extent to which the classification algorithm is able to distinguish between patients that require a stay of less than 2 days or those who require longer. Randomly picking between two outcomes has AUC 0.5; the maximum is 1, indicating perfect discriminating ability.

Our *baseline* is the logistic regression model derived in [1], as we aim to improve upon the results they achieved for trauma patients. All AUC figures are obtained from *10-fold stratified cross-validation*.

7. Results

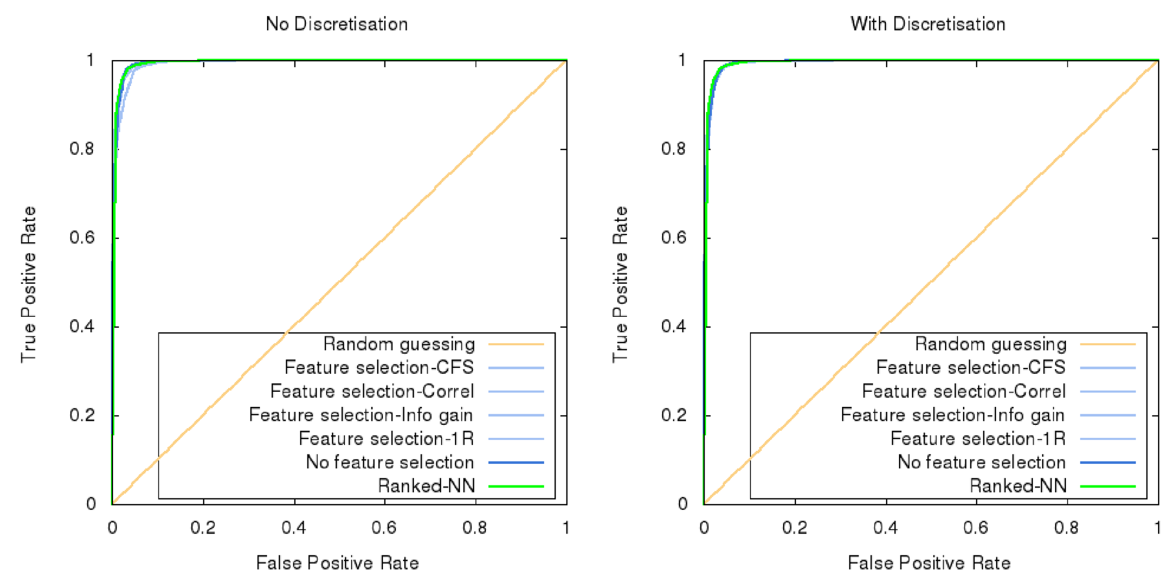
7.1. Trauma dataset

Discretisation significantly improved the performance of all classifiers, regardless of feature selection method. Feature selection usually improved performance but depended on the method, with CFS the most effective. Ranked NN is not affected by discretisation and does not require feature selection, and improves upon baseline AUC by 1-2% (0.82-0.83) while being faster to train than more sophisticated methods. The best classifiers were logistic regression and SVM with discretisation with mean AUC 0.84 (baseline 0.81).



7.2. General dataset

All classification algorithms performed very well regardless of discretisation or feature selection as the original features were very good at predicting LOS. Again, performing discretisation makes the method of feature selection less important, as it stabilises performance overall. The best results were mean AUC 0.994 with logistic regression and SVM regardless of discretisation; ranked NN had mean AUC 0.992 with less training time.



8. Conclusions and Future Work

Careful application of discretisation and feature selection can improve the predictive power of classification algorithms. We proposed a general method to improve the LOS predictions for a trauma and general hospital domain, and showed overall improvements in AUC from 2-4% from the baseline. This general method will be a useful starting point for those investigating LOS prediction.

Furthermore, ranked NN performs 1-2% better than the baseline, is moderately fast to train, and is not strongly affected by discretisation and irrelevant features.

There is room to further evaluate this method in other medical areas, as well as implementation of such a model as a decision support tool for physicians. Additionally, improving upon the ranked NN algorithm to account for different methods of weighting features and calculating ranks can also be investigated.

9. Acknowledgements

Many thanks to Dr Michael Dinh, co-director of Royal Prince Alfred Hospital Trauma Services, for providing the trauma data and medical domain guidance; and thanks also to Associate Professor Paulo Cortez, from the Department of Information Systems at the University of Minho in Portugal, for providing an extensive general LOS dataset.

References

[1] M. M. Dinh, K. J. Bein, C. M. Byrne, B. Gabbe, and R. Ivers. Deriving a prediction rule for short stay admission in trauma patients admitted at a major trauma centre in Australia. *Emergency Medicine Journal*, 0:1–5, 2013.