Using Machine Learning to Refer Patients with Chronic Kidney Disease to Secondary Care

Au-Yeung L[1], Xie X[2], Chess J[3,4], Scale T M[3,4]

[1] Swansea University Medical School, [2] Department of Computer Science, Swansea University, [3] Wales Kidney Research Unit, [4] Swansea Bay Local Health Board,

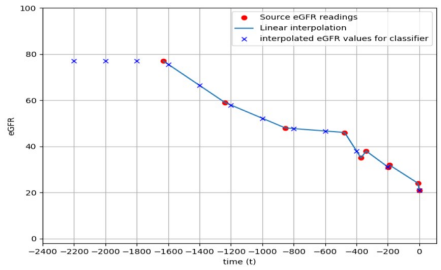
1 Abstract

Machine learning offers a range of classification algorithms which can be applied to medical data to aid in making clinical predictions. Recent studies have demonstrated the high predictive accuracy of various classification algorithms applied to clinical data. Several studies have already been conducted in diagnosing or predicting chronic kidney disease at various stages using different sets of variables. In this study we are investigating the use of machine learning techniques with blood test data. Such a system could aid renal teams in making recommendations to primary care general practitioners to refer patients to secondary care where patients may benefit from earlier specialist assessment and medical intervention.

2 Background

A typical graph of eGFR (estimated Glomerular Filtration Rate) readings over time for a patient presented by the Assist-CKD application to operators.

3 Readings at Regular Intervals



Imputing values using linear interpolation between 2 nearest points.

4 Construct Feature Matrix

Feature Matrix

	imputed eGFR_1	...	imputed eGFR_n	Age	Sex
Patient_1	55	:	30	89	1
Patient_2	54	:	25	32	0
Patient_3	53	:	45	48	1
:	:	:	:	:	:
Patient_n	:	:	:	:	:

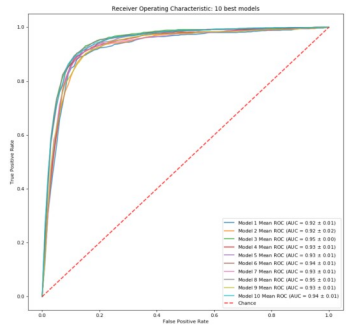
Quantitative Evaluation

Model #	Avg Training Time (s)	Avg Overall Accuracy	Avg Sensitivity	Avg Specificity
LogReg, L, A, S	6.47	88.48%	86.67%	89.02%
LogReg, L, S	6.71	88.14%	86.50%	88.63%
LogReg, L, A	6.42	88.09%	86.03%	88.71%
LogReg, L	6.5	88.05%	86.08%	88.64%
ANN (512,64,2), L, S	48.21	87.12%	88.36%	86.74%
ANN (1024,256,2), L	31.93	86.94%	89.01%	86.31%
ANN (512,64,2), L	15.21	86.60%	89.74%	85.65%
ANN (1024,256,2),L,A	15.69	86.61%	89.30%	85.81%
SVM(LK), L, A	0.4	85.29%	85.51%	85.23%
SVM(LK), L, A, S	0.39	84.78%	80.29%	86.14%

Key: L—Late Aligned, A—Age, S—Sex, LK-Linear Kernel, ANN-Artificial Neural Network, SVM-Support Vector Machine

ROC Curves

Comparative ROC cuves for top 10 performing classification models.



Results

We are able to achieve an overall accuracy of 88.48% using logistic regression, 87.12% using ANN and 85.29% using SVM. ANNs performed with the highest sensitivity at 89.74% compared to 86.67% for logistic regression and 85.51% for SVM.

Conclusions

Based on the results of our experimentation we would recommend using a logisitic regression based classifier initially since they performed with the highest overall accuracy. Artificial Neural Networks (ANN) merit serious consideration because they performed with the highest sensitivity, which is very desirable in a clinical context.

Future Work

It would be interesting to explore the use of convolutional neural networks to provide a more automated method of feature extraction.

Acknowledgements

Special Thanks to Dr James Chess and Dr Timothy Scale, Wales Kidney Research Unit.

References

[1] K. J. et. al., "Prevalence and management of chronic kidney disease in primary care patients in the uk," https://www.ncbi.nlm.nih.gov/pubmed/24852335, 2014.

[2] H. P. et. al., "Diagnosis of chronic kidney disease based on support vector machine by feature selection methods," https://www.ncbi.nlm.nih.gov/pubmed/28243816, 2017.

[3] M. D. et. al., "Patient classification and outcome prediction in iga nephropathy," https://www.ncbi.nlm.nih.gov/pubmed/26453758, 2015.

[4] J. N. et. al., "Predicting renal failure progression in chronic kidney disease using integrated intelligent fuzzy expert system," https://www.hindawi.com/journals/cmmm/2016/6080814/abs/, 2016.

[5] K. J. et. al., "Machine-learning approaches to assist in accurate and extensive chronic kidney disease screening," http://www.ijcea.com/machine-learning-approaches-assist-accurateextensive-chronic-kidney-disease-screening/, September 2017.

[6] C. M. Bishop, "Pattern recognition and machine learning," 2006.

[7] B. Z. Riccardo Bellazzi, "Predictive data mining in clinical medicine: Current issues and guidelines," https://www.ncbi.nlm.nih.gov/pubmed/17188928, 2008.

[8] T. D. N. et. al., "An end stage kidney disease predictor based on an artificial neural networks ensemble," https://www.sciencedirect.com/science/article/pii/S0957417413000778, 2013.

[9] W. G. Baxt, "Application of artificial neural networks to clinical medicine," https://www.ncbi.nlm.nih.gov/pubmed/7475607, 1995.

[10] R. P. Lippmann, "An introduction to computing with neural nets," https://ieeexplore.ieee.org/abstract/document/1165576, 1987, [Online; accessed 28/09/2019].