Galles, Oscar; Caravaca Rodríguez, Miriam; Suppi Boldrito, Remo; [et al.]. «Prediction of chronic kidney disease progression with artificial intelligence : a challenge within our reach». Nephrology Dialysis Transplantation, Vol. 38, Issue Supplement$_1(June 2023), art. gfad063c_4640. DOI 10.1093/ndt/gfad063c_4640$

PREDICTION OF CHRONIC KIDNEY DISEASE PROGRESSSION WITH ARTIFICIAL INTELLIGENCE: A CHALLENGE WITHIN OUR REACH

Oscar Galles[1], Miriam Caravaca Rodríguez[1], Remo Suppi[2], Edwar Macias[3], Antoni Morell[4], Jordi Comas[5], Elisenda Martinez[6], Tomas Salas[6] and Jose Ibeas[1,7]

[1]Institut of Research and Innovation of Parc Taulí University Hospital, Clinical, Interventional and Computational Nephrology, Sabadell, Barcelona, Spain, [2]School of Engineering. Universitat Autònoma de Barcelona, Computer Architecture and Operating Systems, Barcelona, Spain, [3]School of Engineering. Universitat Autònoma de Barcelona, Barcelona, Spain, [4]School of Engineering. Universitat Autònoma de Barcelona, Telecommunications and Systems Engineering, Barcelona, Spain, [5]Catalan Transplantation Organisation (OCATT). Health Deparment., Catalan Renal Registry (RMRC), Barcelona, Spain, [6]Health Quality and Assessment Agency of Catalonia (AQuAS), Barcelona, Spain and [7]Parc Taulí University Hospital, Universitat Autònoma de Barcelona, Nephrology, Sabadell, Barcelona, Spain

BACKGROUND AND AIMS:
Chronic Kidney Disease (CKD) is a common and debilitating condition that affects over 850 million people worldwide. The disease is associated with high mortality rates that can reach up to 10-15% per year, multiple complications, among which cardiovascular ones stand out. These complications can contribute to the progression of CKD, and this in turn to the appearance of complications, feeding each other. Despite the availability of clinical guidelines and predictive models, accurately predicting disease progression and identifying risk factors for progression in CKD patients remains a challenge. The limitations of current methodologies, including simplifying complex relationships and relying on linear assumptions, have hindered progress in this area. The advancement of Artificial Intelligence and Machine Learning has provided a new opportunity to address these challenges. The goal of this study was to evaluate the performance of gradient boosting algorithms in predicting the progression of renal disease in a large dataset of 1327 patients with a follow up of 10 years.

METHOD:
Design: Retrospective analysis of a historical cohort from the Register of Renal Patients of Catalonia (RMRC) and the Data analytics program for health research and innovation (PADRIS) from Health Quality and Assessment Agency of Catalonia (AQuAS). Inclusion Criteria: > 18 y.o. CKD stages from 2 to Renal Replacement Therapy (RRT) and adequate data after pre-processing the sample. N = 1.327 patients with 27.572 records. Follow up of 10 years (January 2010 - December 2020). Variables: Age, gender, BMI, Diagnoses (ICD-10) = 95, Transplant waiting list status; RRT status; Laboratory variables = 77; Pharmacological treatment = 100.
Method:
By using Light Gradient-Boosting Machine (LGBM) testing CKD progression prediction horizon in quarterly windows for multiple periods.
Methodology: 1. Pre-processing of the sample and data. 2. Training and testing for variables exploration. 3. Dataset structuring in quarterly windows. 4. Samples randomization and data separation for a 5-fold cross-validation (20% test - 80% training). 5. Training and tuning of LGBM model for different prediction horizons.

RESULTS:
Age: 62 ± 13 years; Gender: 34% female, 66% male. Best prediction horizon was for 8 quarters (2 years), with a ROC curve of 0.967 and accuracy of 0.860. The 10 variables with major

relevance in the model in order were estimated Glomerular Filtration Rate, Age, Microalbuminuria, BMI, HDL, Glucose, Urea, Platelets, Triglycerides and Sodium.

CONCLUSION:

1. The prediction of CKD progression can benefit from the use of Machine Learning with results that outperform methods based on classical statistics. 2. It can allow the individualization of the prognosis and thus be able to carry out early interventions to improve the prognosis.

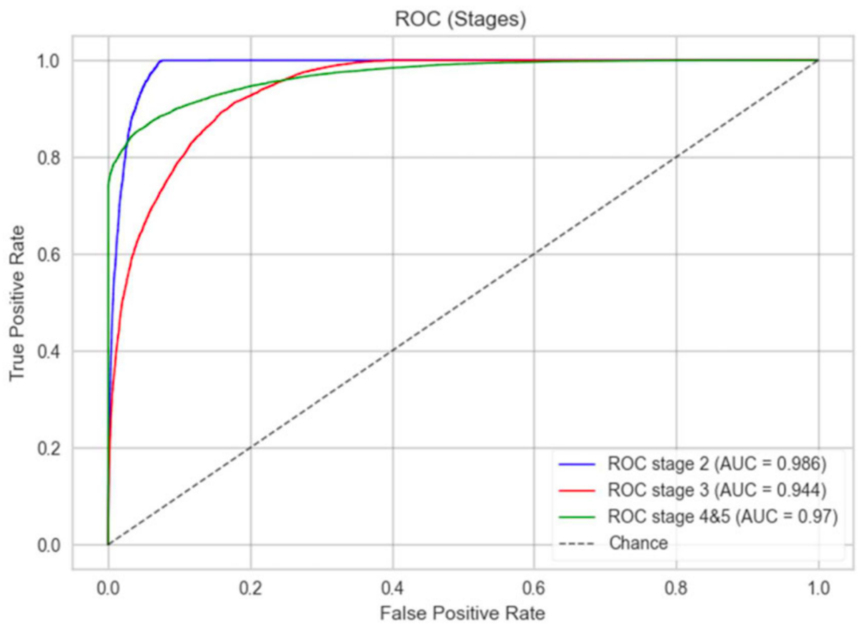

Figure 1: ROC Curve for stages 2, 3, 4, and 5 of CKD.

Stage/Metric	Accuracy	Precision	Recall	F1 Score	AUC
2	-	0.795	0.854	0.824	0.986
3	-	0.793	0.806	0.799	0.944
4 & 5	-	0.920	0.896	0.908	0.970
Avg.	0.860	0.836	0.852	0.844	0.967

Table 1: Metrics list for LGBM Classifier.