APPENDIX FOR THE ECIS 2024 PAPER:

ARE OUR PREDICTIONS HEALTHY? A COMPARATIVE META-ANALYSIS OF MACHINE LEARNING STUDIES IN PREDICTIVE HEALTHCARE

Appendix

Kai Heinrich, OVGU Magdeburg, Germany, kai.heinrich@ovgu.de

Armin Keshavarzi, OVGU Magdeburg, Germany, armin.keshavarzi@st.ovgu.de

Abstract

Predictive healthcare in the case of pancreatic neuroendocrine tumors (PNETs) is a crucial operation as treatment challenges arise due to the heterogeneity of the disease. Surgical approaches vary based on aggressiveness, ranging from resection for milder cases to extensive removal for aggressive PNETs. Thus, machine learning (ML) models are crucial for precise prediction and categorizing PNETs for enhanced outcome forecasting. This systematic review sheds light on the practices of ML approaches within a comparative meta-analysis and a quality assessment employing the standardized IJMEDI checklist. The results show that ML studies within the field of predictive healthcare, despite their potential, face challenges like inadequate data preprocessing, unclear model architecture, and limited clinical applicability.

Keywords: Predictive healthcare, systematic review, meta-analysis, quality.

Author	ML model	Validation method	AUROC (95% CI)	Accuracy	Sensitivity	Specificity
		Aggressiveness				
(J. Huang et al. 2022)	MLR - Clinical DL - Clinical DL - combined	70 - 30 random split	0.78 0.81 0.85	0.67 0.79 0.75	0.58 0.75 0.75	0.75 0.83 0.75
(Javed et al. 2022)	MLR RF	80 - 20 random split	0.72 0.86	0.71 0.82	0.5 0.75	0.78 0.8
(Mori et al. 2022)	RF for Metastasis RF for Grade RF - Radiomics for Metastasis RF - Radiomics for Grade RF - Radiomics for Lymphnodes RF - Radiomics for Microvascular Invasion	70 - 30 random split	0.697 0.717 0.769 0.806 0.689 0.75			
(X.-T. Huang et al. 2022a)	Nomogram scoring RPA	external validation	0.82 0.81	0.75 0.74	0.71 0.7	0.76 0.75
(Yu et al. 2022)	MLR	external validation	0.84	0.8	0.77	0.81

(Zhu et al. 2022)	MLR	35 - 65 random split	0.849	0.86	0.75	
		Diagnosis				
(Thiis-Evensen et al. 2022)	BT* SVM* LDA*	3-flod cross validation	0.98 0.97 0.96	0.89 0.87 0.83	0.95 0.74 0.60	0.727 0.734
		Grades				
(Chiti et al. 2022)	Arterial - Model not reported Venous - Model not reported	75 - 25 random split	0.82 0.6813			
(Liu et al. 2022)	LDA - Clinical LDA - MRI LDA - CT LDA - combined	70 - 30 random split	0.77 0.83 0.75 0.85	0.76 0.83 0.71 0.83	0.72 0.8 0 0.84	
(Otto et al. 2023)	softmax MLR Deconvolution model	80 - 20 random split		0.85 0.81	0.85 0.8	0.96 0.97
(Park et al. 2023)	ANN - Clinical - Grades ANN - Radiomics - Grades ANN - combined - Grades RF - Clinical - Grades RF - Radiomics - Grades RF- combined - Grades ANN - Clinical - prognosis ANN - Radiomics - prognosis ANN - combined - prognosis RF - Clinical - prognosis RF - Radiomics - prognosis RF- combined- prognosis	5-fold cross validation	0.705 0.857 0.864 0.664 0.819 0.853 0.728 0.662 0.83 0.72 0.596 0.741	0.655 0.724 0.776 0.603 0.751 0.828 0.672 0.655 0.776 0.724 0.569 0.707	0.68 0.64 0.72 0.48 0.64 0.8 0.741 0.667 0.776 0.741 0.481 0.593	0.89 0.95 0.99 0.69 0.58 0.58 0.63 0.881 0.636 0.788 0.818 0.697
(Wang et al. 2022)	SVM-linear - group 1 SVM-linear - group 2 SVM-linear - group 3	60 - 40 random split 5-fold cross	0.84 0.87 0.88	0.75 0.75 0.78	0.83 0.83 0.86	0.79 0.89
		Mortality				
(Jiang et al. 2023)	Cox proportional hazards Neural Multitask Logistic Regression DeepSurv Random Survival Forest	70 - 30 random split 5-fold cross validation	0.87 0.87 0.9 0.86	0.74 0.84 NaN NaN	0.7501 0.7616 0.7882 0.7612	0.1397 0.1418 0.1278 0.1432
(Liao et al. 2022)	Cox models, RF	70 - 30 random split	Nan	NaN	0.76	
(Lu et al. 2022)	Cox models, Bayesian network	70 - 30 random split	Nan	NaN	0.82	
(G. Xu et al. 2022)	Cox models	bootstrapping and external validation	0.822	Nan	0.826	
		Recurrence				
(An et al. 2022)	Clinical data - regression model Radiomics - regression model Combined Radiomics - regression model	70 - 30 random split	0.786 0.712 0.824			
(Murakami et al. 2023)	RSF - 1 year RSF - 5 year RSF - 10 year Cox models 1 year Cox models 5 year	70 - 30 random split	0.937 0.835 0.911 0.936 0.737	0.97 0.98 0.86 0.97 0.98	0.841 0.841 0.841 0.82 0.82	0.108 0.108 0.108 0.151 0.151

Cox models 10 year	0.81	0.85	0.82	0.151

Table A1. Machine Learning models properties
Abbreviations: MLR: Multivariable logistic regression, RF: Random Forest, ANN: Artificial Neural Network, DL: Deep learning model, RPA: recursion partitioning analysis, BT: Boosted Tree, SVM: Support vector machines, LDA: Linear discriminant analysis, RSF: Random Survival Forest, *: multiple models were averaged

Category	Output	Model (with best AUROC)	AUORC	Predictive Features	Author
Aggressiveness	Aggressive and Non-aggressive prediction	DL - combined	0.85	Sex, Age, Functional, CEA, CA125, CA 19–9, Tumor location, Body, Tail, Tumor size, Texture, Tumor shape, Tumor margin, Echogenicity, DMPD, CDFI, AE, VE, ED	(J. Huang et al. 2022)
	Nodal disease prediction in nonfunctional PNET	RF	0.86	Tumor size, Tumor location, arterial enhancement, portal venous enhancement, and 15 radiomic features	(Javed et al. 2022)
	Grades, Metastasis, Lymph nodes, Microvascular aggressiveness	RF	0.806	Age, Gender, Necrosis, Cystic morphology, Pancreas atrophy, Arterial invasion, Venous invasion, Contiguous organs invasion, Grade (G1 vs. G2/G3), Liver metastasis (M+), Microvascular invasion (VI+), Metastatic lymph nodes (N+)	(Mori et al. 2022)
	Lymph node metastasis (LNM) prediction	Nomogram scoring	0.82	Age, sex, tumor size, tumor location, serum CgA level, serum NSE level, and Ki-67	(X.-T. Huang et al. 2022a)
	Ki-67 index prediction of being less than 5%	MLR	0.84	Tumor size, arterial phase enhancement, portal venous phase enhancement, and arterial phase enhancement pattern	(Yu et al. 2022)
	lymph node metastasis prediction	MLR	0.849	Gender, Age, BMI, Symptom, NLR, TB, ALT, AST, FBG, CEA, CA199, CA724, NSE, Lymph node metastasis, Tumor location, SI on T2WI, Maximum diameter of the tumor, Tumor margin, Exophytic growth, MPDD or CBDD, Hyperenhancement at arterial phase, Homogeneity, Vascular and adjacent tissue involvement, Synchronous liver metastases, Long axis of the largest lymph node, Short axis of the largest lymph node, Ratio of the long/short axis of the largest lymph node, Abnormal Shape of the largest lymph node, Number of the lymph nodes with the short axis>5 mm, Number of the lymph nodes with the	(Zhu et al. 2022)

				short axis>10 mm, ADCmean, ADCmax, ADCmin, Tumor volume	
Diagnosis	Diagnosis prediction	BT*	0.98	Age, sex, Distant metastasis, Lymph node metastasis, Ki-67, NET Grade	(Thiis-Evensen et al. 2022)
Grades	Prognosis prediction	ANN - combined - prognosis	0.83	Tumor grade prediction model: Clinical features (1) Primary tumor size (2) Hepatic metastasis (3) Extrahepatic metastasis (4) Tumor grade Radiomic features (1) Morphological surface to volume ratio (2) Morphological center of mass shift (3) Intensity-based kurtosis (4) Intensity-histogram minimum histogram gradient grey level (5) GLCM correlation (6) GLRLM long runs emphasis (7) NGTDM strength (8) GLSZM small zone emphasis (9) GLSZM large zone emphasis (10) GLSZM large zone low grey level emphasis (11) GLSZM grey level non-uniformity; Prognosis prediction model: Clinical features (1) Sex (2) Clinical stage (3) Vascular invasion Radiomic features (1) Morphological spherical disproportion (2) Morphological sphericity (3) Intensity-histogram kurtosis (4) Intensity-histogram quartile coefficient of dispersion (5) Intensity-histogram maximum histogram gradient (6) GLSZM normalized grey level non-uniformity	(Park et al. 2023)
	Grades classes	ANN - combined - - Grades	0.864		
Mortality	Survival rate prediction	DeepSurv	0.9	Age, Gender, Marital status, Race, Primary site, Stage, Grade, Surgery, Radiotherapy, Chemotherapy, Tumor size, Number of tumors, Tumor extension, Distant metastasis, Survival months, Status	(Jiang et al. 2023)
Recurrence	Prognosis prediction of GP-NENs	Combined Radiomics - regression	0.824	Gender, History of hypertension, Smoking history, Drinking history, Age, Tumor pathological type, Primary tumor	(An et al. 2022)

		model		site, Ki-67, TNM stage, Lymph node metastasis, Distant metastasis, History of diabetes, Radscore 1, Radscore 2, Radscore 3	
	1-year recurrence prediction	RSF	0.937	Age, Sex, Hereditary syndrome, Symptom, Tumor location, Tumor number, Contrast pattern, Cystic component, Calcification, Main pancreatic duct obstruction, Surgical method, Surgical procedure, Lymphadenectomy, Clavien-Dindo classification, Tumor grade, Tumor size, Lymph node metastasis, Stage, Residual tumor, Lymphovascular invasion, Perineural invasion	(Murakami et al. 2023)
	5-year recurrence prediction	RSF	0.835		
	10-year recurrence prediction	RSF	0.911		

Table A2. summary of the predictive features and models with highest AUROC among each category and similar ourput.