Journal of Gastrointestinal Cancer
https://doi.org/10.1007/s12029-023-00960-1

REVIEW

Non-endoscopic Applications of Machine Learning in Gastric Cancer: A Systematic Review

Marianne Linley L. Sy-Janairo[1] · **Jose Isagani B. Janairo**[2]

Accepted: 16 July 2023

Abstract

Purpose Gastric cancer is an important health burden characterized by high prevalence and mortality rate. Upper gastrointestinal endoscopy coupled with biopsy is the primary means in which gastric cancer is diagnosed, and most of machine learning (ML) tools are developed in this area. This systematic review focuses on the applications of ML in gastric cancer that do not involve endoscopic image recognition.

Methods A systematic review of ML applications that do not involve endoscopy and are relevant to gastric cancer was performed in two databases and independently evaluated by the two authors. Information collected from the included studies are year of publication, ML algorithm, ML performance, specimen used to create the ML model, and clinical application of the model.

Results From 791 screened studies, 63 studies were included in the systematic review. The included studies demonstrate that the non-endoscopic applications of ML can be divided into three main categories, which are diagnostics, predicting response to therapy, and prognosis prediction. Various specimen and algorithms were found to be used for these applications. Most of its clinical use includes histopathologic slide reading in the diagnosis of gastric cancer and a risk scoring system to determine the survival of patients or to determine the important variables that may affect the patient's prognosis.

Conclusion The systematic review suggests that there are numerous examples of non-endoscopic applications of ML that are relevant to gastric cancer. These studies have utilized various specimens, even non-conventional ones, thus showing great promise for the development of more non-invasive techniques. However, most of these studies are still in the early stages and will take more time before they can be clinically deployed. Moving forward, researchers in this field of study are encouraged to improve data curation and annotation, improve model interpretability, and compare model performance with the currently accepted standard in the clinical practice.

Keywords Machine learning · Gastric cancer · Artificial intelligence

Introduction

Gastric cancer is one of the most common cancers worldwide. According to the Global Cancer Observatory (GLOBOCAN) data from the World Health Organization (WHO), it is the 5th most frequently diagnosed cancer. Despite its decreasing incidence in the past years, it remains to be an important health burden due to its high prevalence and mortality rate [1]. It accounts for 5.6% of total cancer incidence and 7.7% of cancer-related deaths in 2020 for both sexes across all ages [2]. GC patients often present with abdominal pain and unintentional weight loss [3]. Evaluation and diagnosis of patients include blood tests, biopsy, upper endoscopy, and imaging studies. Treatment depends on the preoperative staging and includes endoscopic resection, surgical resection, chemotherapy, immunotherapy, and radiotherapy. In the United States, most patients with gastric cancer are symptomatic and have advanced disease at the time of presentation [4]. Late-stage GC patients have poor prognosis, which is why it is critically important to devise strategies and interventions that can lead to earlier detection which can enhance the survival rate for the disease.

✉ Jose Isagani B. Janairo
jose.isagani.janairo@dlsu.edu.ph

[1] Institute of Digestive and Liver Diseases, St. Luke's Medical Center–Global City, Taguig, Philippines

[2] Department of Biology, De La Salle University, 2401 Taft Avenue, 0922 Manila, Philippines

Published online: 21 July 2023

The rapid advancement of machine learning (ML) and artificial intelligence (AI) has catalyzed the development of innovative technologies relevant to medicine. ML and AI refer to a range of computational tools that learn from data to predict a specific outcome. ML models can derive insights or identify patterns from data, enabling them to make predictions. The workflow in creating ML models includes data collection and cleaning. The data can be split into training, validation, and test sets, wherein the training set is used to "teach" the predictive model, the validation set to tune and optimize the parameters, and the test set to assess the performance of the optimized model. External validation dataset can also be used to further test the performance of the ML model since the external validation dataset contains data that was not part of the training, validation, and test sets. Thus, external validation provides a glimpse of the real-world performance of the model. The model "learns" through an iterative process, characterized by repetitively adjusting the parameters based on the training data. The initial model starts with randomly assigned values for the different parameters and makes predictions on the training data. The difference between the predicted and actual values is monitored by a loss function, which quantifies the difference between them. By applying optimization algorithms, the model modifies its parameters to minimize the loss. This leads to an improvement of the predictions as the cycle is repeated. This process continues for multiple iterations, gradually refining the model's ability to make accurate predictions based on patterns in the training data. The inner workings of ML models and a detailed discussion on how different algorithms work are beyond the scope of this systematic review. There is a rich and accessible literature about the details on how ML works [5, 6], which interested readers can later on pursue.

The improved collection and storage of clinical data coupled with the popularization of data science may have led the rapid development of ML tools for a wide array of applications. In medicine, one of the most visible ML applications is the commercially available AI-assisted endoscopy. The tool aims to enhance polyp detection by scanning the field of view of the endoscope in real time for the presence of polyps. Despite the technology is still in its infancy, this tool exhibits a lot of promise. Available evidence suggests that incorporating AI in endoscopy enhances polyp detection [7]. In gastric cancer management, AI has proven to be effective in enhancing the detection of early stages of gastric cancer based on endoscopic images [8]. It is evident that machine vision—the processing and analysis of image data using AI in endoscopy—is maturing faster compared with other ML applications. This systematic review will thus focus on the emerging applications of AI that do not involve endoscopic image recognition. This review aims to present recent ML developments that have the potential to improve the diagnosis and management of gastric cancer.

Methodology

Literature Search

Two authors independently conducted a search for articles published in January 2012 to August 2022 in PUBMED (538 articles) and Scopus database (253 articles). The keywords used were "artificial intelligence" OR "machine learning," AND "gastric cancer" OR "gastric malignancy," AND NOT "Endoscopic." Different institutions from different countries were included to allow generalizability of the study. All retrieved abstracts were independently reviewed by the two authors to determine eligible articles. Disagreements were discussed by the two authors until the issue is resolved and a consensus is reached. 115 articles on full texts were reviewed for validity and eligibility criteria. Figure 1 shows the flow diagram of the search study.

Study Selection

All retrieved abstracts were independently reviewed by two authors. The full texts of potential articles were retrieved and reviewed by the same authors to determine eligible articles. Disagreements were discussed by the two authors until the issue is resolved and a consensus is reached. Studies that involved the use of machine learning in the clinical applications of gastric cancer were included. Exclusion criteria were as follows: (i) the application of ML is marginal, (ii) studies are non-clinical (ex. dealt with mechanism or carcinogenesis of gastric cancer), (iii) methodology was unclear or did not present complete information, (iv) papers are letters, commentaries, editorials, reviews, systematic review, and meta-analysis, (v) full-text article was not able to be retrieved despite best effort, and (vi) the article was not written in English. The open-source software R [9] was used to create all the diagrams in this paper using the *circlize* package [10].

Results

A total of 69 articles were included in this systematic review. Broadly, the papers can be categorized into three main groups: (1) applications related to the diagnosis, (2) applications related to therapy efficacy, and (3) applications related to prognosis. Studies that involve two or more clinical applications were included in both categories. The first study retrieved in this review was published in 2017. The

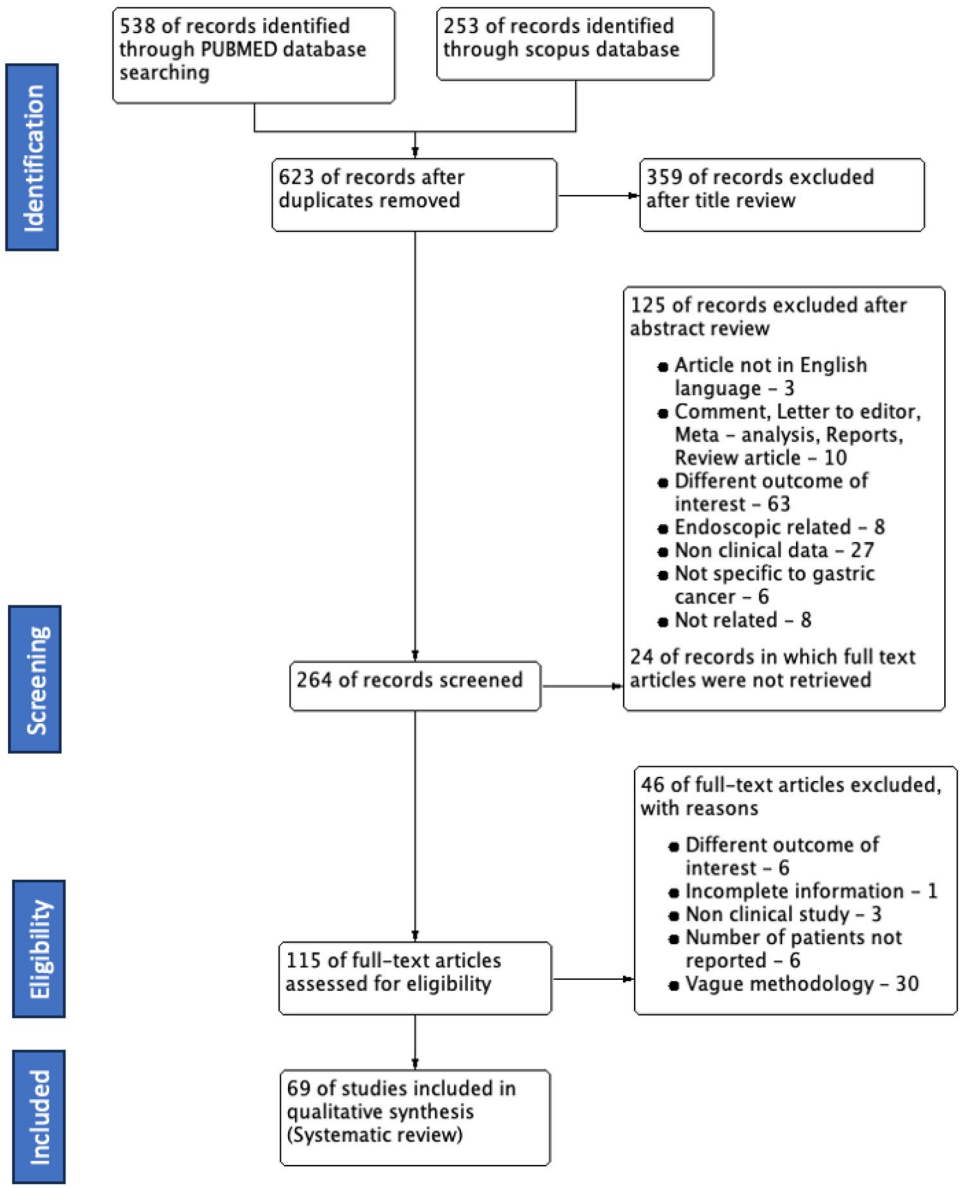

Fig. 1 Flow diagram of the search strategy

largest sample size was 23,067 in the study of Liu et al. in 2022 [11].

Applications Related to Diagnosis

Table 1 summarizes the details of the papers that are relevant in GC diagnosis that were included in this systematic review. The indicated model performance pertains to the performance of the best or optimized model on the test or validation dataset. A total of 25 articles were included in this category. Figure 2 shows the chord diagram visualizing the algorithm–specimen for diagnostics. It is evident from Fig. 2 that deep learning algorithms are the most widely used for this application class, while the most common clinical use involves histopathologic slides. AI-assisted imaging reading, especially with CT scan, is also gaining popularity. Aside from assisting in the diagnosis and staging of gastric cancer, it was also found that machine learning has been utilized in the identification of gastric cancer subtypes and potential biomarkers that may serve as non-invasive screening modality. Saliva and breath characteristics have also been explored as potential way to discriminate GC patients from healthy individuals.

Applications Related to Therapy Efficacy

The details of the papers included in this systematic review relevant to this application are summarized in Table 2. Eighteen articles were included in this category, and most of the papers involve the evaluation or prediction of response

Table 1 Summary of studies that met the inclusion criteria and are relevant to GC diagnostics

Reference	Year	Algorithm	Population	Performance	Specimen	Clinical applications
[12]	2017	SVM	94 patients	AUC = 0.99	Salivary transcriptomic data	The model is used to identify potential salivary biomarkers for the non-invasive detection of gastric cancer
[13]	2017	Deep learning	454 patients	AUC = 0.699	Histopathology slides	The proposed convolutional neural network architecture reports favorable results for cancer classification and detection of necrosis
[14]	2018	SVM	124 patients	AUC = 0.904	Gene expression data	This study identified an activated miRNA module biomarker of GC, which was composed of miR-486, miR-451, miR-185, and miR-600. The hub miRNAs and subnetwork of GC were also detected by a NetSVM-based method. These identified hub miRNAs and modules have the potential to become robust biomarkers for the diagnosis of early-stage GC
[15]	2020	LASSO	539 patients	AUC = 0.758	Clinical and CT scan images	The model uses radiomic signature and CT-derived data to predict the Lauren classification in gastric cancer
[16]	2021	Deep learning	7 patients (16,822 total images)	AUC = 0.9986	Histopathology slides	The model can diagnose diffuse gastric cancer based on images of slide specimens
[17]	2021	Linear discriminant analysis	36 patients	AUC = 0.928	Blood	Blood plasma phospholipids were used as features for the discrimination of GC patients from healthy individuals
[18]	2021	Naïve Bayes	80 patients	AUC = 0.9	MRI images	The Bayesian model can categorize the stage of the gastric tumor using MRI-derived data
[19]	2021	Deep learning	200 patients	Acc = 0.895	Breath	The model can detect gastric cancer using breath features derived from Raman spectroscopy
[20]	2021	Random forest	213 subjects	AUC = 0.9199	Blood	Raman spectroscopy signals of the collected serum samples were used as features for the classifier that can distinguish serum in GC patients from healthy individuals
[21]	2021	SVM	279 subjects	AUC = 0.9984	Gene expression data	The model can discriminate gastric adenocarcinoma from healthy controls

Table 1 (continued)

Reference	Year	Algorithm	Population	Performance	Specimen	Clinical applications
[22]	2021	Deep learning	373 patients	AUC = 0.907	Histopathology slides	A virtual cytokeratin staining and deep learning-based tumor-stroma ratio measurement is used to aid in the diagnosis of TSR in gastric cancer
[23]	2021	Deep learning	375 patients	AUC = 0.85	Histopathology slides	The deep learning-based algorithm is used to directly predict EBV infection in gastric cancer from H&E-stained histopathology slides
[24]	2021	SVM	539 subjects	AUC = 0.84	CT scan images	It was found that 2D features of CT images provided better predictive performance to identify the type of GC
[25]	2021	Deep learning	831 patients	AUC = 0.947	Histopathology slides	The model can identify gastric cancer subtypes based on pathology images
[26]	2021	Deep learning	1164 patients	AUC = 0.990	Histopathology slides	The model is used for accurate and efficient identification of metastatic lymph node
[27]	2021	Deep learning	1570 subjects	Acc = 0.91, Sn = 0.98, Sp = 0.85	Tongue images	Using tongue images, the deep learning model can identify GC from healthy individuals
[28]	2021	Deep learning	1978 patients	AUC = 0.946	Clinical, pathologic characteristics, and CT scan images	The model may identify patients with clinically occult peritoneal metastasis using CT scan images and other clinical data
[29]	2021	Deep learning	2823 patients	AUC = 0.863	Histopathology slides	Detection of microsatellite instability and Epstein-Barr virus-associated gastric cancer from histopathology using a deep learning-based classifier
[30]	2021	Deep learning	2824 specimens from distinct patients	AUC = 0.99	Histopathology slides	The model can identify gastric signet ring cell carcinoma in whole slide images
[31]	2021	Random forest	2951 patients	Acc = 0.9138, Sn = 0.9067, Sp = 0.9216, AUC = 0.9720	Blood	The model can differentiate blood samples derived from healthy to GC-inflicted individuals. The model uses 17 clinical blood features, such as monocyte ratio and total protein, among other parameters

Table 1 (continued)

Reference	Year	Algorithm	Population	Performance	Specimen	Clinical applications
[32]	2022	Artificial neural network	62 patients	Acc = 0.85, Sn = 1, Sp = 0.708, AUC = 0.95	Saliva	The model uses salivary markers as potential diagnostic test for GC
[33]	2022	Naïve Bayes	139 subjects	AUC = 0.817	Breath	Breath characteristics in the form of volatile organic compounds were used as features for ML classifier that can identify GC patients from the control group
[34]	2022	SVM	357 patients	AUC = 0.9224	Clinical and CT scan images	The model is used for preoperative evaluation of HER2 status in patients with gastric cancer
[35]	2022	Deep learning	1383 patients	AUC = 0.941	Histopathology slides	The model can identify Epstein-Barr virus-associated gastric cancer from histopathology
[36]	2022	XGBoost	3630 patients	AUC = 0.896	Clinical and lifestyle data	The model is used to predict the risk for gastric cancer using clinical and lifestyle behavior

to chemotherapy. Other use of machine learning in this category involves assessment of preoperative resectability and outcomes of postsurgical intervention. Figure 3 shows the chord diagram visualizing the algorithm–specimen for therapy efficacy. It can be observed from the selected studies that multiple data types are the often-used specimen to build ML models for this application. Multiple data types in this context mean that more than 2 different types of data were used to build the ML model. Various data types in this category include clinical profiles, surgical/pathologic characteristics, and demographics, among others. In addition, it was observed that SVM was the most common algorithm in this application class for model building. Apart from multiple data types, genetic data were also observed as another significant specimen that is used in this category.

Applications Related to Prognosis

Majority of the included studies in this systematic review falls under this category, which comprise 36 articles with their corresponding details shown in Table 3. Application related to prognosis mainly deals with the prediction of the survival of patients with gastric cancer, with or without treatment. Papers that dealt with prediction of lymph node metastasis and recurrence of gastric cancer were also included under this category. The chord diagram showing the relationship between specimen type and algorithm used to build ML models for prognosis prediction applications is seen in Fig. 4. ML applications in this area used a variety of algorithms but mostly fall within the category of tree ensemble which includes decision trees, gradient boosting, and random forests. Similar to applications for predicting response to therapy, multiple data types comprise the main specimen or data source of ML models for prognosis. The models were used to create a risk scoring system to determine the survival of patients or to determine the important variables that may affect the patient's prognosis. Radiologic information has also been widely explored for this category.

Discussion

It was observed in this review that aside from the current use of ML in the field of endoscopy, its main usefulness in the diagnosis of gastric cancer is related to the improvement of histopathologic slide reading. ML models are now being explored to assist in histopathologic slide reading. The primary reason for this is that the gold standard for diagnosing gastric cancer is through biopsy during upper gastrointestinal tract endoscopy. The use of ML in diagnostic pathology may help the pathologist in increasing the detection of gastric cancer and diagnosing its subtype. This is important because it may affect the management

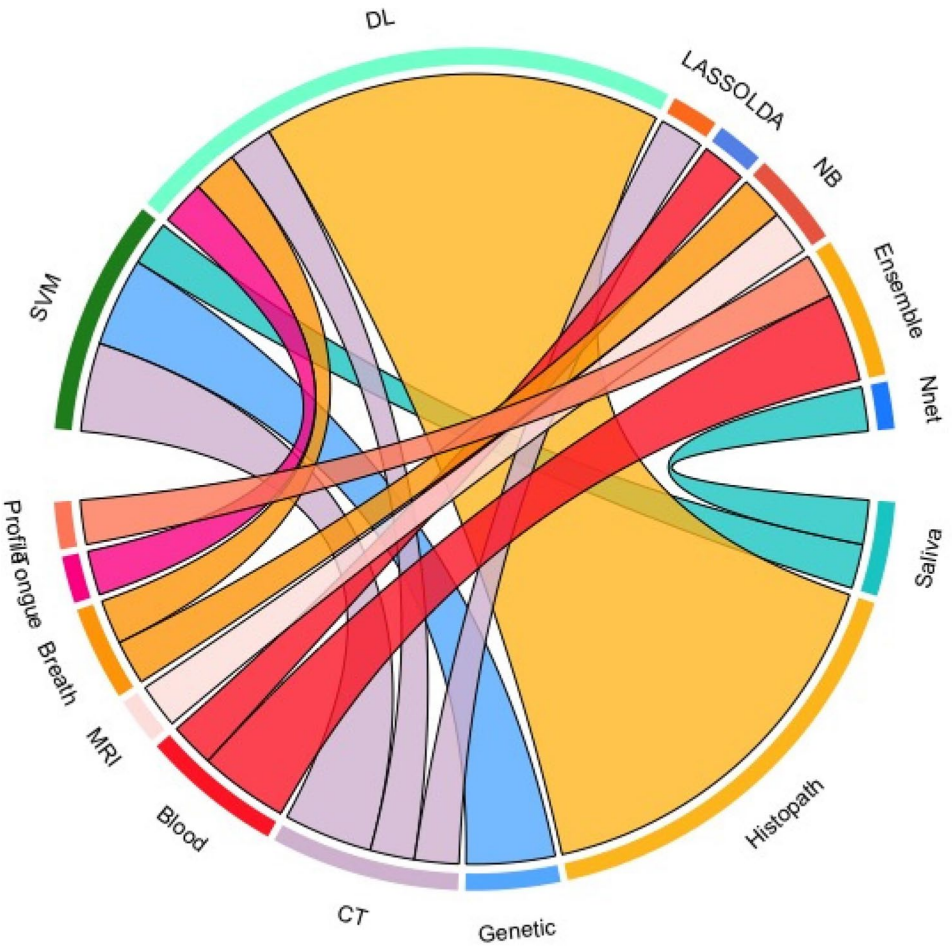

Fig. 2 Chord diagram showing the relationship between specimen type and algorithm used to build ML models for applications relevant to GC diagnostics. SVM = support vector machine, DL = deep learning, LASSO = least absolute shrinkage and selection operator, LDA = linear discriminant analysis, NB = naïve Bayes, Nnet = artificial neural networks

of the patient, as certain subtype may require different treatment. Different gastric cancer subtype may respond differently to certain treatment and carries different prognosis [80]. Furthermore, the advancement in this field may also translate to shorter turnaround time in the histopathology reading and lower the miss out detection rate. In terms of the machine learning algorithms used, it can be observed that deep learning models were often applied to histopathology images. Deep learning is an advanced type of artificial neural network that possesses multiple layers of nodes which enable these models to extract patterns from various types of data, including from images. It is therefore not surprising to note that is widely used for model building in this application class.

Based on the studies included in this systematic review, it can be observed that significant developments have been achieved in the non-endoscopic applications of ML for the diagnosis and management of gastric cancer. It can also be observed that various specimens including non-conventional ones are being used to create these models. This review has showed that ML is being used to explore saliva, breath, and tongue characteristics to discriminate gastric cancer patients from healthy individuals. ML tools are likewise being used to identify potential biomarkers and develop a more sensitive and less invasive screening tests. This is a commendable initiative since this may lead to the development of more non-invasive tests and procedures that are relevant to the early detection of gastric cancer. Less-invasive procedures can encourage more people to undergo screening tests which can aid in the earlier diagnosis of cancers. The early diagnosis of gastric cancer is of paramount importance because it translates to early initiation of treatment and better prognosis. Its five-year survival rate is inversely correlated with the stage of the cancer at the time of diagnosis and management. The five-year survival rate treated with surgical intervention for stage IA tumors is 94% while it is at 18% for stage IIIC tumors [81]. However, early gastric cancer is difficult to diagnose because most patients at this stage are asymptomatic or present with non-specific symptoms. Several screening tests for gastric cancer are available, but still require further improvement. The development of exploratory ML tools to create a more sensitive and specific test to screen and diagnose gastric cancer is a welcome development and may lead to increased detection of gastric cancer at an earlier stage, which may have a significant impact in their prognosis.

Table 2 Summary of studies that met the inclusion criteria and are relevant to predicting response to therapy

Reference	Year	Algorithm	Population	Performance	Specimen	Clinical applications
[37]	2017	SVM	584 patients	AUC = 0.78	Clinical, pathological, and surgical characteristics	The classifier was able to distinguish patients with stage II and III gastric cancer who have significant differences in the 5-year disease-free survival and overall survival with chemotherapy
[38]	2018	Random forest	30 patients	AUC = 0.722	CT scan images	The model can predict the response of patients with chemotherapy using CT images
[39]	2018	SVM	786 patients	AUC = 0.834	Clinical, pathological, proteomics, and surgical characteristics	The developed SVM classifier can accurately predict overall survival and disease-free survival. Moreover, the classifier can identify predict patients with stage II and III GC that can benefit from adjuvant chemotherapy
[40]	2019	SVM	321 patients	Acc = 0.781	Clinical, pathological, and surgical characteristics	Useful predictor for measuring the risk of postoperative morbidities and may help stratify patients with different overall status for choice of surgical procedures or other treatments
[41]	2019	SVM	668 patients	AUC = 0.815	Clinical, pathological, and surgical characteristics	The SVM model is composed of seven features related to tumor microenvironment and can predict survival among patients with stage II and III gastric cancer and who will benefit from treatment
[42]	2021	SVM	124 patients	AUC = 0.854	Clinical, pathological, and CT scan images	The predictive model demonstrated excellent agreement using CECT-based radiomic analysis across three different cancer hospitals, each achieving AUCs of > 0.80 in predicting response to neoadjuvant chemotherapy
[43]	2021	Logistic regression	221 subjects	AUC = 0.744	CT scan images and clinical characteristics	Using segment CT images and other clinical data as the features, the model can classify GC patients based on their response to chemotherapy
[44]	2021	SVM	292 subjects	AUC = 0.889	CT scan images	The model can predict pathological downstaging outcomes using CT images
[45]	2021	Deep learning	1615 patients	AUC = 0.789	Clinical, pathological, surgical, and imaging characteristics	The model can predict disease-free survival using CT images which can help identify patients that will most likely benefit from treatment
[46]	2022	Deep learning	102 subjects	Acc = 0.941, Sn = 0.956, Sp = 0.818	Contrast-enhanced ultrasound and CT scan images	The model can assist in the evaluation of preoperative resectability of GC patients
[47]	2022	SVM	444 subjects	AUC = 0.85	Genetic/gene expression data, clinical, and CT scan images	The model analyzed the gene expression and assessed the tumor microenvironment to predict survival and adjuvant chemotherapy benefit of gastric cancer patients after radical gastrectomy

Table 2 (continued)

Reference	Year	Algorithm	Population	Performance	Specimen	Clinical applications
[48]	2022	Non-negative matrix factorization	567 subjects	AUC = 0.981	Gene expression data	The cluster analysis revealed important genes for prediction of survival and response to adjuvant chemotherapy and immunotherapy after gastrectomy and to immune checkpoint inhibitors for gastric cancer
[49]	2022	Deep learning	719 patients	AUC = 0.827	Clinical and imaging characteristics	The model has satisfactory ability to discriminate positive response to neoadjuvant chemotherapy
[50]	2022	LASSO	729 subjects	AUC = 0.919	Gene expression data	The model can categorize patients based on response to chemotherapy using focal adhesion-related genes
[51]	2022	Boruta feature selection	1304 subjects	Not applicable	Clinical and surgical characteristics	The model was used to identify important variables related to overall survival of GC patients following therapy
[52]	2022	Random forest	1459 subjects	AUC = 0.88	Gene expression data	The model can predict disease-free survival of greater than or less than 2 years for patients on paclitaxel using identified genomic biomarkers
[53]	2022	Logistic regression	Multiple curated datasets were used	AUC = 0.8	Gene expression data	The model can differentiate patients that respond and do not respond to immunotherapy

It is evident in this review that the ML models are also being used to assist imaging reading of gastric cancer, particularly in CT scan. The studies have showed that ML-assisted image reading has potential use in the assessment of preoperative resectability and accurate staging. Aside from it, another application of ML model is predicting the prognosis using certain radiologic features. Radiomics is widely gaining clinical use as it is now being explored in the prediction of subtype, prediction of locoregional metastases, and prediction of recurrence as observed in this review.

Assessing and predicting patient response to therapeutic interventions is important to GC management because it determines the effectiveness of the treatment. Currently, this is achieved through a wait-and-see approach wherein patients are periodically evaluated after the treatment, clinically and radiographically. As newer treatment modalities become available, predicting patient's response becomes important in choosing the appropriate regimen that will provide greater clinical benefit. In this review, it was noted that ML models are built from combination of multiple data types such as clinical profiles, surgical/pathologic characteristics, and demographics, among others, to predict patient's response and effectiveness to treatment. Furthermore, genetic information has been explored to contribute on the ML models for this prediction. This may have a profound impact in the management of gastric cancer as treatment becomes more personalized. These developments can be interpreted as greater efforts to understand the impacts of various treatment regimens on the patient and at the cellular level. Thus, these developments are expected to contribute to tailored treatment strategies based on the genomic information of individuals. These help in the development of a personalized gastric cancer management strategy. It may help select the appropriate regimen in which the patient may benefit the most. ML models can minimize the uncertainty in selecting the best therapeutic options and can therefore potentially improve prognosis. In terms of the type of ML, SVM is widely used to analyze data from data type. SVM is a type of algorithm that can model non-linear data and rely on the creation decision boundaries using hyperplanes to differentiate one categorical variable from the other and is deemed to be appropriate in this category.

It is evident from the review that most of the published paper for non-endoscopic use of ML for gastric cancer involves prediction of patient's prognosis. Predicting patient prognosis is an important part of GC management because it gives certainty on what to expect in the course of the disease. It helps in the decision-making of the patient, family, and entire healthcare team. Furthermore, it may help in identifying clinical factors that may impact the patient's overall survival. The common algorithms in this application class can be generally categorized as ensemble learning. This type of ML algorithm involves the creation of multiple predictive models and uses

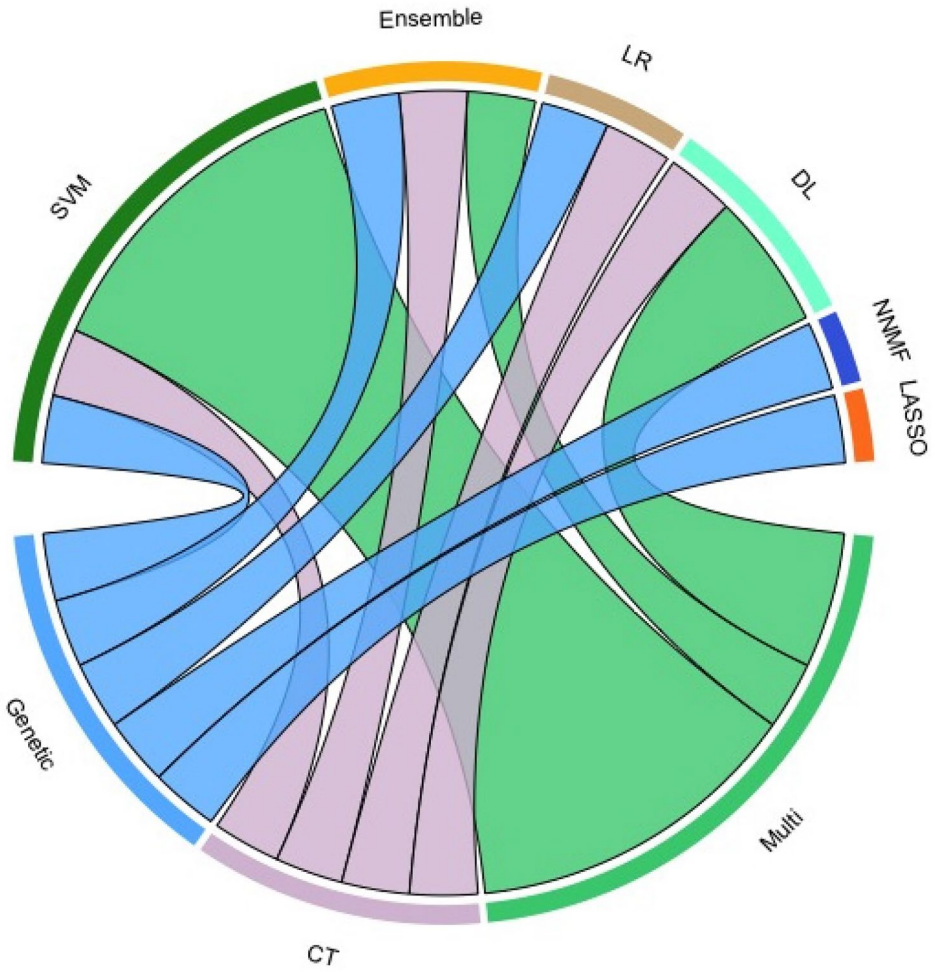

Fig. 3 Chord diagram showing the relationship between specimen type and algorithm used to build ML models for applications relevant to predicting response to therapy. SVM = support vector machine, LR = logistic regression, DL = deep learning, NNMF = non-negative matrix factorization, LASSO = least absolute shrinkage and selection operator

the insights derived from each model to arrive at the final decision or prediction. Due to the ability of ML models to extract relationships from complex data, their deployment to address problems in predicting patient prognosis is highly beneficial to GC management. In this review, most of the applications of ML have been in solving classification problems. In classification models, the ML model categorizes the data into classes using input features. Predicting if a gastric cancer patient will survive after 5 years or not using patient data as the input features is an example of a classification problem.

Measuring the performance of the ML model is a crucial aspect of model creation, wherein the AUC is a common metric for binary classification models. AUC means area under the receiver operating characteristic (ROC) curve and is a measure of the model's ability to differentiate the two classes. A model that possesses an AUC value the falls in between 0.7 and 0.8 is considered to demonstrate acceptable discrimination while an AUC in between 0.8 and 0.9 is considered to demonstrate excellent demonstration. Finally, an AUC value that is greater than 0.9 is considered to have outstanding discrimination ability [82]. Other commonly used performance metrics for classification models are accuracy, precision, recall, $F1$, and Matthews Correlation Coefficient (MCC). These metrics are calculated through the following equations:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$F1 = 2\frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

$$\text{MCC} = \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}}$$

Table 3 Summary of studies that met the inclusion criteria and are relevant to prognosis

Reference	Year	Algorithm	Population	Performance	Specimen	Clinical applications
[37]	2017	SVM	584 patients	AUC = 0.78	Clinical, pathological, and surgical characteristics	The classifier was able to distinguish patients with stage II and III gastric cancer who have significant differences in the 5-year disease-free survival and overall survival with chemotherapy
[54]	2018	Bayesian neural networks	339 patients	AUC = 0.961	Clinical characteristics	The model is used to predict survival of gastric cancer patients
[39]	2018	SVM	786 patients	AUC = 0.834	Clinical, pathological, proteomics, and surgical characteristics	The developed SVM classifier can accurately predict overall survival and disease-free survival. Moreover, the classifier can identify predict patients with stage II and III GC that can benefit from adjuvant chemotherapy
[55]	2018	Recurrent neural network	1243 patients	AUC = 0.81	Clinical, pathological, and surgical characteristics	The ML model can accurately predict survival among GC patients. The model performs better compared with the TNM staging
[41]	2019	SVM	668 patients	AUC = 0.815	Clinical, pathological, and surgical characteristics	The SVM model is composed of seven features related to tumor microenvironment and can predict survival among patients with stage II and III gastric cancer and who will benefit from treatment
[56]	2019	Artificial neural network	1860 patients	AUC = 0.79	Clinical and pathological characteristics	The preoperative model built using ANN can accurately predict the long-term survival of GC patients. Moreover, the performance of the model is not inferior compared to the standard
[57]	2020	Deep learning	204 patients	AUC = 0.76	CT scan images	The deep learning radiomic nomogram showed good performance in predicting LNM in gastric cancer
[58]	2020	LASSO	353 patients	*C*-index = 0.727	Clinical, pathologic characteristics, and CT scan images	An MDCT-guided prognostic model for GC was developed
[59]	2020	Deep learning	730 patients	*C*-index = 0.822	Clinical and CT scan images	A deep learning-based radiomic nomogram had good predictive value for LNM in LAGC
[60]	2020	Artificial neural network	15,959 patients	AUC = 0.789	Clinical and pathological characteristics	The ANN model can predict the survival of GC patients using clinicopathological data
[61]	2021	Logistic regression	159 patients	AUC = 0.908	CT scan images	The model is used for the prediction of lymph node metastasis in early gastric cancer
[62]	2021	Adaboost	185 patients	AUC = 0.822	PET/CT scan images	The model uses the preoperative 18F-FDG PET/CT images to predict the lymph node metastasis and nodal stage of gastric cancer patients
[63]	2021	Cox regression	389 patients	AUC = 0.723	Genetic/gene expression data	The model can predict the survival risk based on microRNA profiles

Table 3 (continued)

Reference	Year	Algorithm	Population	Performance	Specimen	Clinical applications
[64]	2021	Deep learning	840 patients	Not applicable	CT scan images	The machine learning model was used to process and segment CT images, allowing the estimation of body composition. These were then used as features for the estimation of survival of GC patients with sarcopenic obesity
[65]	2021	Cox regression	1061 patients	*C*-index = 0.783	Clinical, pathological, surgical, and imaging characteristics	Deep learning was used to extract features which were used as input variables for the prediction of survival
[26]	2021	Deep learning	1164 patients	AUC = 0.990	Histopathology slides	The model is used for accurate and efficient identification of metastatic lymph node
[66]	2021	XGBoost	1169 subjects	AUC = 0.881	Clinical, pathological, and surgical characteristics	The model can predict lymph node metastasis among GC patients wherein variables related to tumor were found to be important
[45]	2021	Deep learning	1615 patients	AUC = 0.789	Clinical, pathological, surgical, and imaging characteristics	The model can predict disease-free survival using CT images which can help identify patients that will most likely benefit from treatment
[67]	2021	Random forest	2012 patients	AUC = 0.96	Clinical, pathological, and surgical characteristics	The model can predict recurrence of GC after operation
[68]	2021	Random survival forest	2,931 patients	tAUC = 0.80	Clinical, pathological, and surgical characteristics	A large dataset from Wales and England was used to construct a prognostic model for GC surgery patients. 10 variables were identified as important for making the prediction
[69]	2022	Graph neural networks	172 patients	AUC = 0.96	Histopathology slides and immunohistochemistry images	The classification model can categorize patients into short-term or long-term survival
[70]	2022	Random survival forest	289 patients	AUC = 0.613	Clinical characteristics	The model can predict patient survival, with prealbumin level and prothrombin time as important features
[71]	2022	SVM	375 patients	AUC = 0.761	Genetic/gene expression data	The model can predict the survival of gastric cancer patients based on m7G-related genes
[47]	2022	SVM	444 patients	AUC = 0.85	Genetic/gene expression data, clinical, and CT scan images	The model analyzed the gene expression and assessed the tumor microenvironment to predict survival and adjuvant chemotherapy benefit of gastric cancer patients after radical gastrectomy
[72]	2022	SVM	523 patients	AUC = 0.796	Clinical and CT scan images	The predictive performance for lymph node metastasis of all the radiological models was significantly better than that of the clinical model
[73]	2022	Random forest	526 patients	AUC = 0.912	Clinical, pathological, surgical, and imaging characteristics	The prediction model can be used to decide on treatment options and identify lymph node metastasis among patients with undifferentiated type of early gastric cancer

Table 3 (continued)

Reference	Year	Algorithm	Population	Performance	Specimen	Clinical applications
[48]	2022	Non-negative matrix factorization	567 subjects	AUC = 0.981	Genetic/gene expression data	The cluster analysis revealed important genes for prediction of survival and response to adjuvant chemotherapy and immunotherapy after gastrectomy and to immune checkpoint inhibitors for gastric cancer
[74]	2022	Convolutional neural networks	621 subjects	AUC = 0.881	Clinical and surgical characteristics	The model can predict survival or death using 15 clinical features. It was found that hemoglobin was the most important variable in the model
[51]	2022	Boruta feature selection	1304 subjects	Not applicable	Clinical and surgical characteristics	The model was used to identify important variables related to overall survival of GC patients following therapy
[75]	2022	Random survival forest	1360 subjects	AUC = 0.814	Clinical, pathological, and surgical characteristics	The model can predict survival of patients that underwent resection due to gastric cancer
[52]	2022	Random forest	1459 subjects	AUC = 0.88	Genetic/gene expression data	The model can predict disease-free survival of greater than or less than 2 years for patients on paclitaxel using identified genomic biomarkers
[76]	2022	LASSO	2290 subjects	AUC = 0.882	Genetic/gene expression data	A scoring system that can predict survival based on senescence using gene expression data
[77]	2022	Regularized dual averaging	2294 subjects	AUC = 0.73	Clinical, pathological, and surgical characteristics	The model can predict lymph node metastasis among GC patients wherein variables related to tumor size, grade, and depth of invasion were found to be important
[78]	2022	Deep learning	2320 subjects	AUC = 0.843	CT scan images	The model can predict peritoneal recurrence using CT images
[79]	2022	SVM	12,631 patients	AUC = 0.89	Clinical and pathological characteristics	The model can be used to identify lymph node metastasis risk in early gastric cancer after endoscopic resection
[11]	2022	LASSO	23,067 subjects	AUC = 0.81	Clinical, pathological, and surgical characteristics	The model can classify post gastrectomy GC patients into high or moderate-risk group for survival based on clinical data

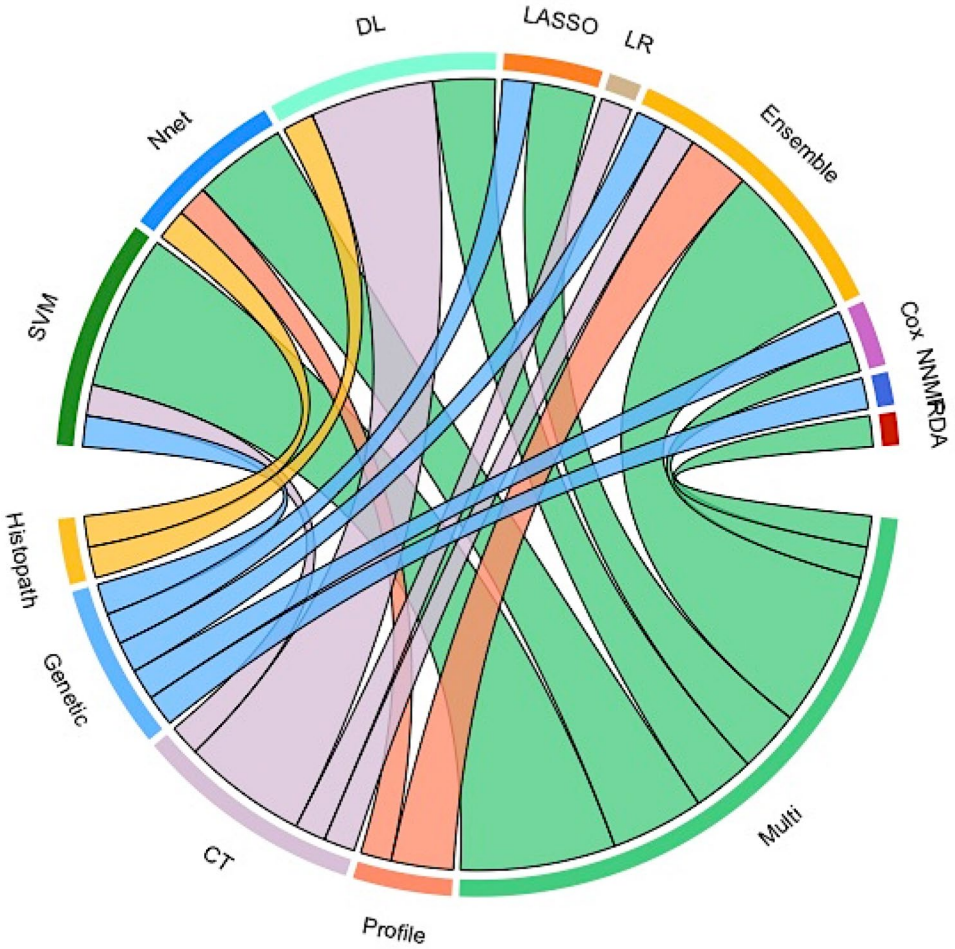

Fig. 4 Chord diagram showing the relationship between specimen type and algorithm used to build ML models for prognosis prediction applications. SVM = support vector machine, Nnet = artificial neural network, DL = deep learning, NNMF = non-negative matrix factorization, LASSO = least absolute shrinkage and selection operator, RDA = regularized dual averaging

where TP = true positive, TN = true negative, FP = false positive, and FN = false negative are derived from the confusion matrix. A confusion matrix summarizes and breaks down the performance of the classifier into four quadrants, each represented by TP, TN, FP, and FN. Various performance metrics are derived from the confusion matrix, as demonstrated by the equations shown above. Accuracy represents the ratio of correctly classified peptides to the total number of peptides. It provides an assessment of overall correctness in classification. The $F1$ score provides a comprehensive view of the classifier's effectiveness as it combines precision and recall into a single metric. It offers insight into both the accuracy of positive predictions (precision) and the ability to identify all positive instances (recall). MCC is a holistic measure of classification performance. It uses the values from all quadrants of the confusion matrix. By considering the overall performance, MCC provides a balanced evaluation of the classifier's ability to handle different types of classifications, which makes it an ideal performance metrics for classification models [83]. For regression models, the predicted variable is continuous, and hence, a different set of performance metrics is used. Mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) are the usual metrics used for assessing regression models and are determined through the following equations:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2$$

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}}$$

where y_i represents the actual value for the response variable and $\hat{y}_i$ is the predicted value. MAE measures the average absolute difference between the predicted and actual values and thus provides a straightforward understanding of the average magnitude of errors. MSE measures the average squared difference between the predicted and actual values. MSE is easily affected by larger errors because of the square operation on the error terms. Finally, RMSE is the square root of the MSE. It provides an interpretable metric

since it has the same units with the target variable. These performance metrics are likewise important in determining if the model overfits. Overfitting occurs when the model parameters become too tailored to the training set which leads to poor generalization on new input data. An indication of overfitting is when the model performs excellently in the training set, but yields poor performance in the test set. Some simple ways of addressing overfitting are by reducing model complexity or the number of variables used to make the prediction, increasing the training population, and through cross-validation.

Compared with endoscopic applications of ML however, it appears that most reported studies are still in the proof-of-concept or preproduction stage wherein the presented models are not yet ready to be deployed in a clinical setting. The endoscopic applications of ML have already been commercialized, such as the EndoBrain which was released by Olympus in 2019 [84]. Thus, the next step in this research landscape is to further develop these models and deploy them in production. In order to realize this goal, certain gaps should be addressed. The first issue that should be addressed is exerting greater effort to ensure that the models are accurate and robust and the findings are generalizable. To achieve this, it is recommended to compare the performance of the model with the currently accepted standard or method. Only a few papers that were included in this systematic review compared the performance of the model with a standard, such as an experienced pathologist. Doing this will provide greater confidence in the capability of the models in their intended application. Another recommendation is exerting greater effort in curating datasets. Since the quality of ML models hinges on the quality and quantity of the data used for training and validation, using high population and well-represented datasets can lead to more reliable models. In addition, the proper labeling of the data points is likewise needed to be checked. During the conduct of this review, we came across studies that used image features as input data for the model, wherein multiple images were taken from each patient. While this practice is totally acceptable, it should be ensured that little to no overlap exists between the training and testing data in order to prevent data leakage which can artificially inflate model performance [85]. To build trust in the crafted ML models, methods relevant to labeling, data annotation, and steps carried out to prevent leakage must be explicitly stated. In addition, greater efforts must be exerted as well to improve the interpretability of the model. Being able to clearly articulate the decision-making process of the model will not only build trust in the ML models but will also foster reproducibility. Since these models are intended to be used in a clinical setting wherein the health and well-being of patients are involved, being able to understand how the predictions are reached is of paramount importance. Some suggestions to improve model interpretability include working with inherently interpretable algorithms, such as linear and logistic regression, decision trees, rough sets, and the recently developed enhanced hyperbox models [86, 87]. If the nature of the data cannot be modeled by these interpretable algorithms, then the use of techniques that aim to interpret the model can be used. Some of these techniques include local interpretable model-agnostic explanations (LIME), Shapley Values, and Shapley Additive Explanations (SHAP), among others [88]. Finally, in order to develop clinically relevant and deployable ML models, physicians and ML scientists should also work with other stakeholders, such as patients, user experience designers in order to create valuable models that meet the demands and needs of the users.

Conclusion

It is clear that the application of ML to GC management goes beyond endoscopic assistance. The review demonstrates that the non-endoscopic applications of ML can be divided into three main categories, which are diagnostics, predicting response to therapy, and prognosis prediction. All of these aspects are important in GC management, and the ML models show clinical promise. Through the systematic review, it was cataloged that these ML models utilize a wide array of specimen, including non-conventional ones. This suggests greater interest to better understand the disease and develop less-invasive tests which are important to encourage more people to undergo tests. Moving forward, it is important to ensure that the created models are accurate, robust, and clinically relevant. Thus, greater effort must be exerted in curating and annotating datasets and creating interpretable ML models. In the future, it will be interesting to see how advanced ML models, such as large language models—the technology behind ChatGPT, can impact clinical practice especially to GC management.

Author Contributions Both authors conceptualized and wrote the paper. Both authors contributed equally to the paper.

Declarations

Competing Interests The authors declare no competing interests.

Animal and Human Studies Not applicable.

Conflict of Interest The authors declare no conflict of interests.

References

1. Petryszyn P, Chapelle N, Matysiak-Budnik T. Gastric cancer: where are we heading? Dig. Dis. S. Karger AG; 2020;280–5.
2. Global Cancer Observatory. Stomach . [cited 2022 Dec 22]. https://gco.iarc.fr/today/data/factsheets/cancers/7-Stomach-fact-sheet.pdf.

3. Wanebo HJ, Kennedy BJ, Chmiel J, Steele G, Winchester D, Osteen R. Cancer of the stomach a patient care study by the American College of Surgeons summary background data. Ann. Surg. 1993.
4. Mukkamalla SK, Recio-Boiles A, Babiker H. Gastric cancer. Treasure Island (FL): StatPearls Publishing; 2022.
5. Cohen S. Chapter 2 - the basics of machine learning: strategies and techniques. In: Cohen S, editor. Artif Intell Deep Learn Pathol. Elsevier. 2021;13–40. https://www.sciencedirect.com/science/article/pii/B9780323675383000026.
6. Nakaura T, Higaki T, Awai K, Ikeda O, Yamashita Y. A primer for understanding radiology articles about machine learning and deep learning. Diagn Interv Imaging. 2020;101:765–770. https://www.sciencedirect.com/science/article/pii/S2211568420302461.
7. Hassan C, Spadaccini M, Iannone A, Maselli R, Jovani M, Thoguluva Chandrasekar V, et al. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. Gastrointest Endosc. 2021;93:77–85. https://www.giejournal.org.
8. Chen PC, Lu YR, Kang YN, Chang CC. The accuracy of artificial intelligence in the endoscopic diagnosis of early gastric cancer: pooled analysis study. J Med Internet Res. JMIR Publications Inc.; 2022;24.
9. R Core Team. R: a language and environment for statistical computing. [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2021. Available from: https://www.r-project.org/.
10. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize implements and enhances circular visualization in R. Bioinformatics [Internet]. 2014;30:2811–2. Available from: https://doi.org/10.1093/bioinformatics/btu393.
11. Liu D, Wang X, Li L, Jiang Q, Li X, Liu M, et al. Machine learning-based model for the prognosis of postoperative gastric cancer. Cancer Manag Res. Dove Medical Press Ltd; 2022;14:135–55.
12. Xu Y-G, Cheng M, Zhang X, Sun S-H, BI W-M. Mutual information network-based support vector machinestrategy identifies salivary biomarkers in gastric cancer. J BUON [Internet]. 2017;22:119–25. Available from: http://www.ebi.
13. Sharma H, Zerbe N, Klempert I, Hellwich O, Hufnagl P. Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology. Comput Med Imaging Graph [Internet]. 2017;61:2–13. Available from: https://www.sciencedirect.com/science/article/pii/S0895611117300502.
14. Zhang F, Xu W, Liu J, Liu X, Huo B, Li B, et al. Optimizing miRNA-module diagnostic biomarkers of gastric carcinoma via integrated network analysis. PLoS One [Internet]. Public Library of Science; 2018;13:e0198445-. Available from: https://doi.org/10.1371/journal.pone.0198445.
15. Wang X-X, Ding Y, Wang S-W, Dong D, Li H-L, Chen J, et al. Intratumoral and peritumoral radiomics analysis for preoperative Lauren classification in gastric cancer. Cancer Imaging [Internet]. 2020;20:83. Available from: https://doi.org/10.1186/s40644-020-00358-3.
16. Rasmussen SA, Arnason T, Huang WY. Deep learning for computer-assisted diagnosis of hereditary diffuse gastric cancer. J Pathol Transl Med. Seoul National University; 2021;55:118–24.
17. Saito R, Yoshimura K, Shoda K, Furuya S, Akaike H, Kawaguchi Y, et al. Diagnostic significance of plasma lipid markers and machine learning-based algorithm for gastric cancer. Oncol Lett. Spandidos Publications; 2021;21.
18. Qiao X, Li Z, Li L, Ji C, Li H, Shi T, et al. Preoperative T2-weighted MR imaging texture analysis of gastric cancer: prediction of TNM stages. Abdom Radiol [Internet]. 2021;46:1487–97. Available from: https://doi.org/10.1007/s00261-020-02802-1.
19. Aslam MA, Xue C, Chen Y, Zhang A, Liu M, Wang K, et al. Breath analysis based early gastric cancer classification from deep stacked sparse autoencoder neural network. Sci Rep [Internet]. 2021;11:4014. Available from: https://doi.org/10.1038/s41598-021-83184-2.
20. Li M, He H, Huang G, Lin B, Tian H, Xia K, et al. A novel and rapid serum detection technology for non-invasive screening of gastric cancer based on Raman spectroscopy combined with different machine learning methods. Front Oncol. Frontiers Media S.A.; 2021;11.
21. Dong C, Rao N, Du W, Gao F, Lv X, Wang G, et al. mRBioM: an algorithm for the identification of potential mRNA biomarkers from complete transcriptomic profiles of gastric adenocarcinoma. Front Genet. Frontiers Media S.A.; 2021;12.
22. Hong Y, Heo YJ, Kim B, Lee D, Ahn S, Ha SY, et al. Deep learning-based virtual cytokeratin staining of gastric carcinomas to measure tumor–stroma ratio. Sci Rep [Internet]. 2021;11:19255. Available from: https://doi.org/10.1038/s41598-021-98857-1.
23. Zhang B, Yao K, Xu M, Wu J, Cheng C. Deep learning predicts EBV status in gastric cancer based on spatial patterns of lymphocyte infiltration. Cancers (Basel) [Internet]. 2021;13. Available from: https://www.mdpi.com/2072-6694/13/23/6002.
24. Meng L, Dong D, Chen X, Fang M, Wang R, Li J, et al. 2D and 3D CT radiomic features performance comparison in characterization of gastric cancer: a multi-center study. IEEE J Biomed Heal Informatics. Institute of Electrical and Electronics Engineers Inc.; 2021;25:755–63.
25. Hinata M, Ushiku T. Detecting immunotherapy-sensitive subtype in gastric cancer using histologic image-based deep learning. Sci Rep. Nature Research; 2021;11.
26. Wang X, Chen Y, Gao Y, Zhang H, Guan Z, Dong Z, et al. Predicting gastric cancer outcome from resected lymph node histopathology images using deep learning. Nat Commun [Internet]. 2021;12:1637. Available from: https://doi.org/10.1038/s41467-021-21674-7.
27. Gholami E, Kamel Tabbakh SR, kheirabadi M. Increasing the accuracy in the diagnosis of stomach cancer based on color and lint features of tongue. Biomed Signal Process Control. Elsevier Ltd; 2021;69.
28. Jiang Y, Liang X, Wang W, Chen C, Yuan Q, Zhang X, et al. Noninvasive prediction of occult peritoneal metastasis in gastric cancer using deep learning. JAMA Netw Open [Internet]. 2021;4:e2032269–e2032269. Available from: https://doi.org/10.1001/jamanetworkopen.2020.32269.
29. Muti HS, Heij LR, Keller G, Kohlruss M, Langer R, Dislich B, et al. Development and validation of deep learning classifiers to detect Epstein-Barr virus and microsatellite instability status in gastric cancer: a retrospective multicentre cohort study. Lancet Digit Heal [Internet]. Elsevier; 2021;3:e654–64. Available from: https://doi.org/10.1016/S2589-7500(21)00133-3.
30. Kanavati F, Ichihara S, Rambeau M, Iizuka O, Arihiro K, Tsuneki M. Deep learning models for gastric signet ring cell carcinoma classification in whole slide images. Technol Cancer Res Treat. SAGE Publications Inc.; 2021;20.
31. Wu J, Yang Y, Cheng L, Wu J, Xi L, Ma Y, et al. GCdiscrimination: identification of gastric cancer based on a milliliter of blood. Brief Bioinform Oxford University Press. 2021;22:536–44.
32. Koopaie M, Ghafourian M, Manifar S, Younespour S, Davoudi M, Kolahdooz S, et al. Evaluation of CSTB and DMBT1 expression in saliva of gastric cancer patients and controls. BMC Cancer. BioMed Central Ltd; 2022;22.
33. Polaka I, Bhandari MP, Mezmale L, Anarkulova L, Veliks V, Sivins A, et al. Modular point-of-care breath analyzer and shape taxonomy-based machine learning for gastric cancer detection. Diagnostics. MDPI; 2022;12.
34. Guan X, Lu N, Zhang J. Evaluation of epidermal growth factor receptor 2 status in gastric cancer by CT-based deep learning radiomics nomogram. Front Oncol [Internet]. 2022;12. Available

from: https://www.frontiersin.org/articles/10.3389/fonc.2022.905203.
35. Zheng X, Wang R, Zhang X, Sun Y, Zhang H, Zhao Z, et al. A deep learning model and human-machine fusion for prediction of EBV-associated gastric cancer from histopathology. Nat Commun. Nature Research; 2022;13.
36. Jiang S, Gao H, He J, Shi J, Tong Y, Wu J. Machine learning: a non-invasive prediction method for gastric cancer based on a survey of lifestyle behaviors. Front Artif Intell [Internet]. 2022;5. Available from: https://www.frontiersin.org/articles/10.3389/frai.2022.956385.
37. Jiang Y, Liu W, Li T, Hu Y, Chen S, Xi S, et al. Prognostic and predictive value of p21-activated kinase 6 associated support vector machine classifier in gastric cancer treated by 5-fluorouracil/Oxaliplatin Chemotherapy. eBioMedicine [Internet]. Elsevier; 2017;22:78–88. Available from: https://doi.org/10.1016/j.ebiom.2017.06.028.
38. Li Z, Zhang D, Dai Y, Dong J, Wu L, Li Y, et al. Computed tomography-based radiomics for prediction of neoadjuvant chemotherapy outcomes in locally advanced gastric cancer: a pilot study. Chinese J Cancer Res [Internet]. 2018;30:406–14. Available from: http://www.cjcrcn.org/article/html_9816.html.
39. Jiang Y, Xie J, Han Z, Liu W, Xi S, Huang L, et al. Immunomarker support vector machine classifier for prediction of gastric cancer survival and adjuvant chemotherapeutic benefit. Clin Cancer Res [Internet]. 2018;24:5574–84. Available from: https://doi.org/10.1158/1078-0432.CCR-18-0848.
40. Lu S, Yan M, Li C, Yan C, Zhu Z, Lu W. Machine-learning-assisted prediction of surgical outcomes in patients undergoing gastrectomy. Chinese J Cancer Res [Internet]. 2019;31:797–805. Available from: https://doi.org/10.21147/j.issn.1000-9604.2019.05.09?viewType=HTML.
41. Jiang Y, Xie J, Huang W, Chen H, Xi S, Han Z, et al. Tumor immune microenvironment and chemosensitivity signature for predicting response to chemotherapy in gastric cancer. Cancer Immunol Res [Internet]. 2019;7:2065–73. Available from: https://doi.org/10.1158/2326-6066.CIR-19-0311.
42. Xie K, Cui Y, Zhang D, He W, He Y, Gao D, et al. Pretreatment contrast-enhanced computed tomography radiomics for prediction of pathological regression following neoadjuvant chemotherapy in locally advanced gastric cancer: a preliminary multicenter study. Front Oncol [Internet]. 2022;11. Available from: https://www.frontiersin.org/articles/10.3389/fonc.2021.770758.
43. Chen Y, Wei K, Liu D, Xiang J, Wang G, Meng X, et al. A machine learning model for predicting a major response to neoadjuvant chemotherapy in advanced gastric cancer. Front Oncol. Frontiers Media S.A.; 2021;11.
44. Xu Q, Sun Z, Li X, Ye C, Zhou C, Zhang L, et al. Advanced gastric cancer: CT radiomics prediction and early detection of downstaging with neoadjuvant chemotherapy. Eur Radiol [Internet]. 2021;31:8765–74. Available from: https://doi.org/10.1007/s00330-021-07962-2.
45. Jiang Y, Jin C, Yu H, Wu J, Chen C, Yuan Q, et al. Development and validation of a deep learning CT signature to predict survival and chemotherapy benefit in gastric cancer: a multicenter, retrospective study. Ann Surg [Internet]. 2021;274. Available from: https://journals.lww.com/annalsofsurgery/Fulltext/2021/12000/Development_and_Validation_of_a_Deep_Learning_CT.415.aspx.
46. Xu D, Liu R, Xu H, Zhang Z, Li W, Zhang Y, et al. Adoption of two-dimensional ultrasound gastrointestinal filling contrast on artificial intelligence algorithm in clinical diagnosis of gastric cancer. Comput Math Methods Med. Hindawi Limited; 2022;2022.
47. Chen T, Li X, Mao Q, Wang Y, Li H, Wang C, et al. An artificial intelligence method to assess the tumor microenvironment with treatment outcomes for gastric cancer patients after gastrectomy. J Transl Med. BioMed Central Ltd; 2022;20.
48. Cheong JH, Wang SC, Park S, Porembka MR, Christie AL, Kim H, et al. Development and validation of a prognostic and predictive 32-gene signature for gastric cancer. Nat Commun. Nature Research; 2022;13.
49. Cui Y, Zhang J, Li Z, Wei K, Lei Y, Ren J, et al. A CT-based deep learning radiomics nomogram for predicting the response to neoadjuvant chemotherapy in patients with locally advanced gastric cancer: a multicenter cohort study. eClinicalMedicine [Internet]. Elsevier; 2022;46. Available from: https://doi.org/10.1016/j.eclinm.2022.101348.
50. Tang X, Wu X, Guo T, Jia F, Hu Y, Xing X, et al. Focal adhesion-related signatures predict the treatment efficacy of chemotherapy and prognosis in patients with gastric cancer. Front Oncol. Frontiers Media S.A.; 2022;12.
51. Terranova N, French J, Dai H, Wiens M, Khandelwal A, Ruiz-Garcia A, et al. Pharmacometric modeling and machine learning analyses of prognostic and predictive factors in the JAVELIN Gastric 100 phase III trial of avelumab. CPT Pharmacometrics Syst Pharmacol. American Society for Clinical Pharmacology and Therapeutics; 2022;11:333–47.
52. Sundar R, Barr Kumarakulasinghe N, Huak Chan Y, Yoshida K, Yoshikawa T, Miyagi Y, et al. Machine-learning model derived gene signature predictive of paclitaxel survival benefit in gastric cancer: results from the randomised phase III SAMIT trial. Gut BMJ Publishing Group. 2022;71:676–85.
53. Kong JH, Ha D, Lee J, Kim I, Park M, Im SH, et al. Network-based machine learning approach to predict immunotherapy response in cancer patients. Nat Commun. Nature Res. 2022;13.
54. korhani Kangi A, Bahrampour A. Predicting the survival of gastric cancer patients using artificial and Bayesian neural networks. Asian Pacific J Cancer Prev [Internet]. 2018;19:487–90. Available from: https://journal.waocp.org/article_55156.html.
55. Oh SE, Seo SW, Choi M-G, Sohn TS, Bae JM, Kim S. Prediction of overall survival and novel classification of patients with gastric cancer using the survival recurrent network. Ann Surg Oncol [Internet]. 2018;25:1153–9. Available from: https://doi.org/10.1245/s10434-018-6343-7.
56. Que SJ, Chen QY, Zhong Q, Liu ZY, Wang J Bin, Lin JX, et al. Application of preoperative artificial neural network based on blood biomarkers and clinicopathological parameters for predicting long-term survival of patients with gastric cancer. World J Gastroenterol. Baishideng Publishing Group Co; 2019;25:6451–64.
57. Li J, Dong D, Fang M, Wang R, Tian J, Li H, et al. Dual-energy CT–based deep learning radiomics can improve lymph node metastasis risk prediction for gastric cancer. Eur Radiol [Internet]. 2020;30:2324–33. Available from: https://doi.org/10.1007/s00330-019-06621-x.
58. Wang S, Feng C, Dong D, Li H, Zhou J, Ye Y, et al. Preoperative computed tomography-guided disease-free survival prediction in gastric cancer: a multicenter radiomics study. Med Phys [Internet]. John Wiley & Sons, Ltd; 2020;47:4862–71. Available from: https://doi.org/10.1002/mp.14350.
59. Dong D, Fang M-J, Tang L, Shan X-H, Gao J-B, Giganti F, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study. Ann Oncol [Internet]. Elsevier; 2020;31:912–20. Available from: https://doi.org/10.1016/j.annonc.2020.04.003.
60. Li Z, Wu X, Gao X, Shan F, Ying X, Zhang Y, et al. Development and validation of an artificial neural network prognostic model after gastrectomy for gastric carcinoma: an international multicenter cohort study. Cancer Med [Internet]. John Wiley & Sons,

Ltd; 2020;9:6205–15. Available from: https://doi.org/10.1002/cam4.3245.
61. Wang X, Li C, Fang M, Zhang L, Zhong L, Dong D, et al. Integrating No.3 lymph nodes and primary tumor radiomics to predict lymph node metastasis in T1–2 gastric cancer. BMC Med Imaging [Internet]. 2021;21:58. Available from: https://doi.org/10.1186/s12880-021-00587-3.
62. Liu Q, Li J, Xin B, Sun Y, Feng D, Fulham MJ, et al. 18F-FDG PET/CT radiomics for preoperative prediction of lymph node metastases and nodal staging in gastric cancer. Front Oncol. Frontiers Media S.A.; 2021;11.
63. Xu J, Wen J, Li S, Shen X, You T, Huang Y, et al. Immune-related nine-microRNA signature for predicting the prognosis of gastric cancer. Front Genet. Frontiers Media S.A.; 2021;12.
64. Kim J, Han SH, Kim H Il. Detection of sarcopenic obesity and prediction of long-term survival in patients with gastric cancer using preoperative computed tomography and machine learning. J Surg Oncol. John Wiley and Sons Inc; 2021;124:1347–55.
65. Hao D, Li Q, Feng Q-X, Qi L, Liu X-S, Arefan D, et al. Identifying prognostic markers from clinical, radiomics, and deep learning imaging features for gastric cancer survival prediction. Front Oncol [Internet]. 2022;11. Available from: https://www.frontiersin.org/articles/10.3389/fonc.2021.725889.
66. Zhou CM, Wang Y, Ye HT, Yan S, Ji M, Liu P, et al. Machine learning predicts lymph node metastasis of poorly differentiated-type intramucosal gastric cancer. Sci Rep. Nature Res. 2021;11.
67. Zhou C, Hu J, Wang Y, Ji MH, Tong J, Yang JJ, et al. A machine learning-based predictor for the identification of the recurrence of patients with gastric cancer after operation. Sci Rep. Nature Res. 2021;11.
68. Rahman SA, Maynard N, Trudgill N, Crosby T, Park M, Wahedally H, et al. Prediction of long-term survival after gastrectomy using random survival forests. Br J Surg NLM (Medline). 2021;108:1341–50.
69. Wang Y, Wang YG, Hu C, Li M, Fan Y, Otter N, et al. Cell graph neural networks enable the precise prediction of patient survival in gastric cancer. npj Precis Oncol. Nature Res. 2022;6.
70. Zhang C, Xie M, Zhang Y, Zhang X, Feng C, Wu Z, et al. Determination of survival of gastric cancer patients with distant lymph node metastasis using prealbumin level and prothrombin time: contour plots based on random survival forest algorithm on high-dimensionality clinical and laboratory datasets. J Gastric Cancer. Korean Gastric Cancer Association; 2022;22:120–34.
71. Li XY, Wang SL, Chen DH, Liu H, You JX, Su LX, et al. Construction and validation of a m7G-related gene-based prognostic model for gastric cancer. Front Oncol. Frontiers Media S.A.; 2022;12.
72. Zhang A, Zhao H, Li F, Liang P, Gao J, Cheng M. Computed tomography-based deep-learning prediction of lymph node metastasis risk in locally advanced gastric cancer. Front Oncol [Internet]. 2022;12. Available from: https://www.frontiersin.org/articles/10.3389/fonc.2022.969707.
73. Wei X, Yan XJ, Guo YY, Zhang J, Wang GR, Fayyaz A, et al. Machine learning-based gray-level co-occurrence matrix signature for predicting lymph node metastasis in undifferentiated-type early gastric cancer. World J Gastroenterol. Baishideng Publishing Group Inc; 2022;28:5338–50.
74. Ko Y, Shin H, Shin J, Hur H, Huh J, Park T, et al. Artificial intelligence mortality prediction model for gastric cancer surgery based on body morphometry, nutritional, and surgical information: feasibility study. Appl Sci. MDPI; 2022;12.
75. Jung JO, Crnovrsanin N, Wirsik NM, Nienhüser H, Peters L, Popp F, et al. Machine learning for optimized individual survival prediction in resectable upper gastrointestinal cancer. J Cancer Res Clin Oncol. Springer Science and Business Media Deutschland GmbH; 2022;
76. Zhou L, Niu Z, Wang Y, Zheng Y, Zhu Y, Wang C, et al. Senescence as a dictator of patient outcomes and therapeutic efficacies in human gastric cancer. Cell Death Discov. Springer Nature; 2022;8.
77. Tian HK, Ning ZK, Zong Z, Liu J, Hu CG, Ying HQ, et al. Application of machine learning algorithms to predict lymph node metastasis in early gastric cancer. Front Med. Frontiers Media S.A.; 2022;8.
78. Jiang Y, Zhang Z, Yuan Q, Wang W, Wang H, Li T, et al. Predicting peritoneal recurrence and disease-free survival from CT images in gastric cancer with multitask deep learning: a retrospective study. Lancet Digit Heal Elsevier Ltd. 2022;4:e340–50.
79. Na J-E, Lee Y-C, Kim T-J, Lee H, Won H-H, Min Y-W, et al. Machine learning model to stratify the risk of lymph node metastasis for early gastric cancer: a single-center cohort study. Cancers (Basel) [Internet]. 2022;14. Available from: https://www.mdpi.com/2072-6694/14/5/1121.
80. Sohn BH, Hwang J-E, Jang H-J, Lee H-S, Oh SC, Shim J-J, et al. Clinical significance of four molecular subtypes of gastric cancer identified by The Cancer Genome Atlas Project. Clin Cancer Res [Internet]. 2017;23:4441–9. Available from: https://doi.org/10.1158/1078-0432.CCR-16-2211.
81. Rawla P, Barsouk A. Epidemiology of gastric cancer: global trends, risk factors and prevention. Prz. Gastroenterol. Termedia Publishing House Ltd.; 2019. p. 26–38.
82. Hosmer D, Lemeshow S. Applied Logistic Regression. 2nd ed. New Jersey: John Wiley & Sons, Inc.; 2000.
83. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics [Internet]. 2020;21:6. Available from: https://doi.org/10.1186/s12864-019-6413-7.
84. Ragunath K. Artificial intelligence in gastrointestinal endoscopy: how intelligent can it get? Lancet Oncol [Internet]. Elsevier; 2019;20:1616–7. Available from: https://doi.org/10.1016/S1470-2045(19)30677-1.
85. Tampu IE, Eklund A, Haj-Hosseini N. Inflation of test accuracy due to data leakage in deep learning-based classification of OCT images. Sci Data [Internet]. 2022;9:580. Available from: https://doi.org/10.1038/s41597-022-01618-6.
86. Tan RR, Aviso KB, Janairo JIB, Promentilla MAB. A hyperbox classifier model for identifying secure carbon dioxide reservoirs. J Clean Prod. 2020;272.
87. Janairo JIB, Aviso KB, Promentilla MAB, Tan RR. Enhanced hyperbox classifier model for nanomaterial discovery. Ai. 2020;1:299–311.
88. Molnar C. Interpretable machine learning [internet]. 2nd ed. 2022 [cited 2023 May 27]. Available from: https://christophm.github.io/interpretable-ml-book.