Building Models to Predict the Diagnosis of Emergency Department Patients using Artificial Neural Networks

Carlos Hernández, Jaime Castillo
Departamento de Procesos Industriales
Universidad Católica de Temuco
Temuco, Chile
carlos.hernandez.zavala@uct.cl, jcastill@uct.cl

Abstract

There are circumstances in which patients require immediate medical care. Usually, hospitals have a specialized unit ready to take care of emergencies such as: strokes, heart attacks, sepsis, severe wounds, and other medical conditions that are also classifies as emergencies. This research compared models based on artificial neural networks (ANN) to predict the diagnosis of patients received at the emergency department (ED) of a public hospital. The research was carried out following a classic 4-phase methodology: analysis, design, development, and validation. During the analysis, patient records collected by ED personnel during 2020 were thoroughly reviewed and preprocessed. During the design, some of the records were selected and divided into several subsets following specific criteria. Each subset was a set of independent variables used as input to build a prediction model. Technical considerations relate to the artificial neural networks (ANN) such as the number of hidden layers and the number of epochs, batch size, size of the training and test set, and size of the validation dataset were also defined at these point. The phases of construction and the validation are carried out entirely using the WEKA 3.9.6. Numerous experiments, trial and error adjustments, and replications were necessary to produce the results shown in this article. For the purposes is the present work around 15,000 records were considered. The complete dataset was divided in two parts: 80% for training and test, and the remaining 20% for validation. The approach to predict the patient diagnose considered the construction of ANNs with of independent variables. During the construction several configurations of ANNs were tried to achieve better results. Only the records allocated for training and test dataset is used during the construction. The performance of the proposed models was measured in terms of the rate of correct predictions made with records allocated for validation. The experimental result revealed that, depending on the number of diagnoses, the values of the target class, the proposed models were able to predict correctly between 55% and 85% of the cases. In conclusion, ANN-based models can help predict ED patients' diagnosis with a reasonable degree of certainty. However, the success of the model depends greatly on the numbers of values that the target class has.

Keywords
Diagnosis Prediction, Artificial Neural Network, Machine Learning, Supervised Leaning, Emergency Department.

1. Introduction
Either in public hospitals or in private clinics, urgent care services are a fundamental pillar of the health system. Emergency departments (ED) have to be prepared and equipped to receive patients who required urgent medical treatments that are usually complex and unscheduled. To satisfy a rather unpredictable flux of incoming patients, an ED have to count with enough supplies, medical staff, and infrastructure.

Being able to deliver an adequate medical care when it is needed is crucial to save lives and to complete the recovery of a patient. In Chile, considering the size of the population there is a relatively high number cases that require urgent care every year. Some studies have shown almost 1,100 annual cases per 1,000 population. Even more, regular care cases are also numerous, being in the order of 19 million per year (Santelices and Santelices, 2017). A study by the Organization for Economic Collaboration and Development (OECD) established that Chile was among the countries with the highest annual rate of urgent care cases reaching up to 571 cases per 1,000 population. The average of OECD countries is 308 cases per 1,000 population (Berchet et al, 2015).

The increasing number of patients requesting urgent care is one of the factor that contributes to the almost permanent overcrowding of the public hospitals and health centers. In some cases, it is simply not possible to admit more patients. This fact increases the waiting time, it prolongs the stay of the patient, and it forces the allocation of urgent bed to regular patients. All of the aforementioned has an impact on the service quality perceived by patients and in the delivery of the treatments. In spite of being matter of study for long time, it has not been developed a methodology or technique to help estimate resources requirements to plan and to support ED operations well in advance (Salway et al, 2017).

By means of applying classical statistical models it is possible to infer associations and, in some cases, even to establish causality based on evidence. However, the size of large databases existing in hospitals and clinics, along with the processing power of the latest microprocessor, and popularity of advanced techniques based in machines learning algorithms offer new alternatives to carry out predictive analysis (Mora, 2022).

In recent years, machine learning has been used widely in diverse areas of medicine because there are processes in healthcare that cannot be analyzed with standard statistical procedures. Therefore, artificial intelligence and machine learning algorithms have gained relevance in the data processing to extract information, to classify data, and to create prediction models (Pedrero et al., 2021).

The use of machine learning (ML) in healthcare areas is a matter of interest for researchers, including the study of infections and contagions (Luz et al., 2020). The implementation of recurrent neural networks to predict situation where temporal patterns in physiological characteristics can help predict the risk of mortality in hospitalized patients (Fengyi et al., 2018).

Studies have shown that the used of models based on ML algorithms to predict waiting time can outperform the results of classical models (Hong et al., 2020). The classification of patients received at health centers can be significantly improved by means of ML-based classifiers (Miles et al., 2020). Furthermore, it has been shown that ML-based models outperformed the moving average by 20% when applied to problem of large queues (Pak et al (2021) helping reduce the number of patients with long waiting time (Pak et al (2021). Other works have implemented ML-based models to deal with the high complexity and randomness of delay and wait time patterns (Curtis et al., 2018).

Finally, it seems necessary to investigate new approaches to help overcome the challenging and dynamic environment of an ED. Taking advantage of advanced data science techniques to extract valuable information from the existing data to support the decision making and consequently, to achieve operational improvements leading better uses of the available resources (supplies, staff, and infrastructure), and to improvement in the service quality.

1.1 Objective
To apply machine learning techniques to predict ED patient diagnosis by means of building predictive models based on artificial neural networks.

2. Literature Review
2.1 Machine learning
Machine learning is usually referred as the branch of artificial intelligence (AI) that uses algorithms to find patterns and to learn from datasets through experience. There several types of machine learning algorithms: supervised, unsupervised, and reinforcement algorithms. In supervised learning, the training is carried out using labelled datasets. This means that the class or the value to be predicted is included in the dataset so it can be used for training. In the case of unsupervised learning, instead, the desired class is not known.

2.2 Classification
In machine learning there several important task: classification, regression, and forecasting. Classification can be understood as the determination of the class, a nominal value, in an unseen dataset using a previously trained model. In a regression problem, instead, the objective is the estimation of a numeric value using independent variable. On the other hand, in a forecasting problem time series are used to predict future values.

2.3 Hold out and cross-validation

Holding out implies the splitting up of a dataset into a set for training and another for testing. The test dataset is employed to assess the performance of the classification model with unseen data. Usually the split up proportion is 80% for training and 20% for testing.

On the other hand, cross-validation is the random split up of a dataset into k folds. During the model building, k-1 folds are employed for training while the left one is used to test model's performance. Training and testing are repeated iteratively k times until all folds have been used for testing (Figure 1). The goal is to minimize the risk of overfitting that can happen when holding out. In the case of cross-validation, each iteration produces different results because the folds for training and for testing have been interchanged. These k results are finally averaged.

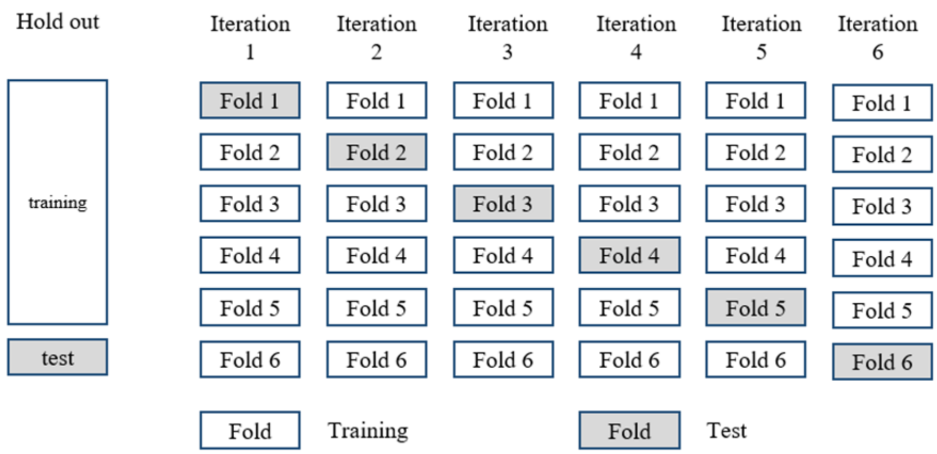

Figure 1. Hold out and cross-validation (k=6)

2.4 Overfitting and generalization

Overfitting occurs when a model learns well from the training dataset but it does not have a good performance when tested with an unseen dataset. In such situation, it is said that the model cannot generalize. This might happen due to the incorporation of many details from the training data that will not be easily found in new data (Figure 2).

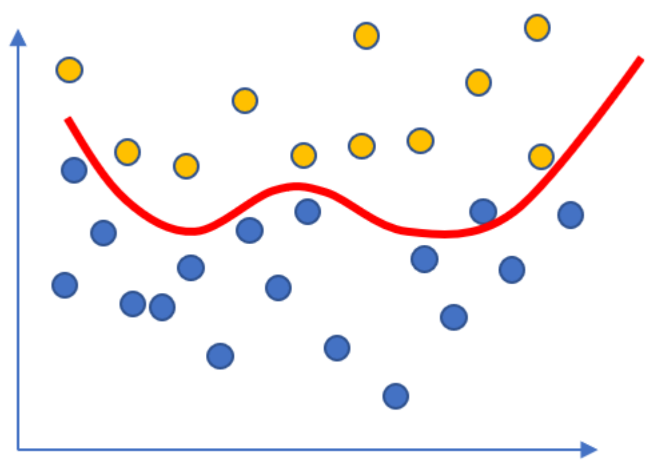

Figure 2. Overfitting

2.5 Replication

Replication is repetition of an experiment under similar conditions to estimate the variability of the results. When using cross-validation, the dataset partitioning into k folds depends on a specific seed number (Figure 3). Since different seed numbers produce different folds, the results of the training and test are different too. By means of replicating the experiments with random seeds each time, it would possible to obtain several test results from which the mean and the standard deviation can be estimated and analyzed afterwards. Thus minimizing the effect of an unfortunate partitioning.

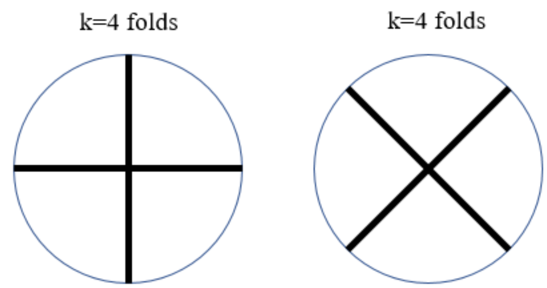

Figure 3. Different folds in cross-validation (k=4)

2.6 Artificial neural network (ANN)

An ANN is a construction made of nodes, referred as neurons, that are combined in an interconnected layered structure (Figure 4). The input layer corresponds to the nodes that receive the externa data. In the second level contains the hidden layers that transform the input data for the output layer, whose neurons are responsible for delivering the results generated by the network.). The topology of an ANN is determined by the number of layers, their nodes, and a transfer function. An interesting exposition on this is given by Morano and Tajani (Morano and Tajani, 2013).

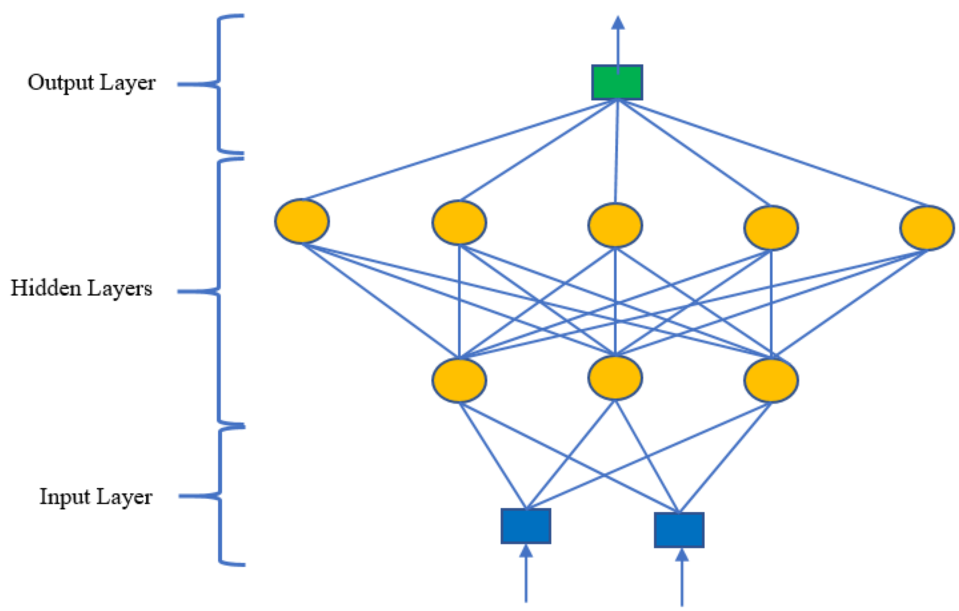

Figure 4. ANN's input, hidden, and output layers

3. Methods

This investigation is carried out following a classic 4-phase model: analysis, design, construction, and validation (Figure 5).

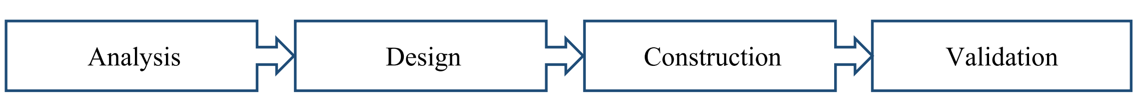

Figure 5. Four-phase model

3.1 Analysis

During the analysis, the preprocessing of the ED database is completed. For the purposes of this investigation, only data corresponding to 2020 were considered. The original dataset comprised a collection of all patient arrival to the ED during a whole year period.

For every patient arrival there is a record. A single record can be understood as a collection of fields. From the arrival time to the payment method. Not all fields were used in this investigation. Instead, only the following were included: age, gender, reason for visit, arrival mode, source of admission, alcohol test, and diagnosis. Being the latter, the target class to be predicted (Table 1).

Table 1. Patient record fields

Field	Description
Age	The age of patient at date of admission.
Gender	The gender of a patient
Reason for visit	The reason(s) for the patient requiring involvement with ED medical staff.
Arrival mode	The principal means by which a patient arrives at ED.
Source of admission	The source of admission to ED.
Alcohol test	Test to detect the presence of alcohol in the patient.
Diagnosis	The identification of the nature of a problem by examination of the symptoms.

An interesting fact revealed during the records' preprocessing is the recurrence of the diagnose. For instance, the five most recurrent diagnose covers almost 14% of all the records in a year with a total 8,423. Twenty-one percent of 2020 records were related to only 10 diagnose. It is important to clarify that each diagnosis has a unique code and that list of codes is rather extensive. In the following, diagnose will be referred by their code (Table 2).

A summary of the number of diagnosis codes and the number of corresponding records is presented in Table 3 and Figure 6. It can be seen that 80% of records corresponds to the most recurrent 10% of diagnose.

Table 2. Record distribution per diagnose

Diagnose	Number of records	Percentage of referrals
5	8.423	14 %
10	12.078	21 %
15	14.304	24 %
20	16.040	27 %

Table 3. Diagnose v/s records

Number of codes	Percentage of total codes	Number of Records	Percentage of total records
327	10 %	47,166	81 %
639	20 %	52,156	89 %
951	30 %	54,402	93 %
1264	40 %	55,702	95 %
1576	50 %	56,528	97 %
1888	60 %	57,152	98 %

1888	60 %	57,152	98 %
2200	70 %	57,505	98 %
2512	80 %	57,817	99 %
2824	90 %	58,129	99 %
3121	100 %	58,426	100 %

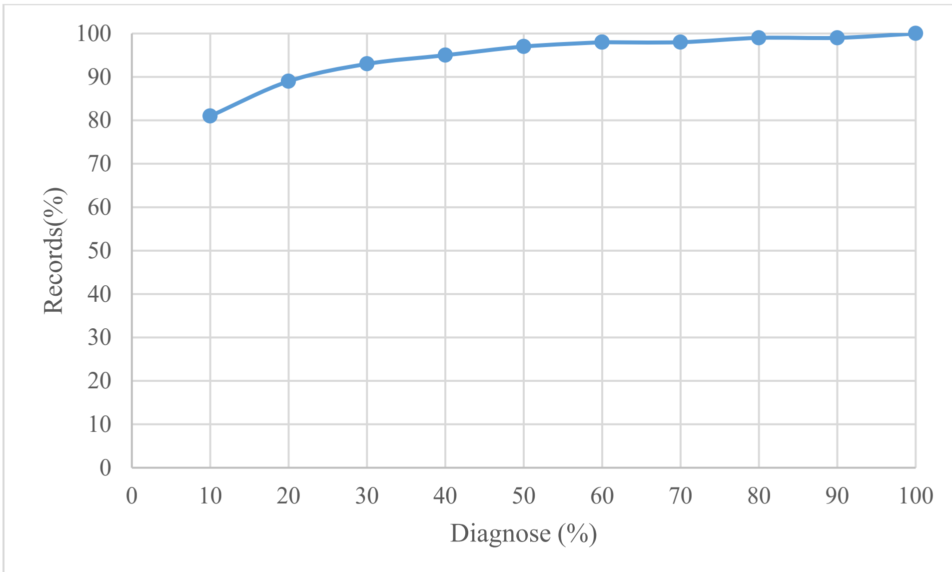

Figure 6. Diagnosis (%) v/s records (%)

3.2 Design

A dataset can be understood as a matrix made of rows that represent patient records or instances, and columns that represent fields or attributes.

During the design phase, four datasets were prepared. The first one containing only the top five most recurrent diagnose. The second one including the top ten diagnose. And the left two one having the 15 and 20 most common diagnose respectively. (Table 4)

Table 4. Dataset creation

Dataset	Number of diagnose	Number of records
DS-1	5	8,423
DS-2	10	12,078
DS-3	15	14,304
DS-4	20	16,040

Each dataset was then split up to create two subsets. The first one for training and test with 80% of data, and the second one for validation with the remaining 20% of data. (Figure 7, Table 5)

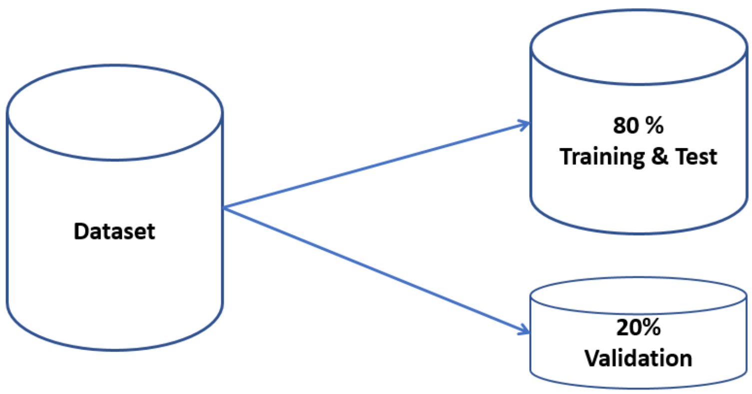

Figure 7. Dataset split up

Table 5. Datasets for training and test, and for validation

Dataset	Records for training and test	Records for validation
DS-1	6,723	1,700
DS-2	9,678	2,400
DS-3	11,404	2,900
DS-4	12,840	3,200

The intention behind crating separated files is to confirm whether the prediction model developed and trained with the training and test set is able to generalized properly when it is used to predict with unknown data from the validation set. The answer to this question will be presented in the latter sections.

As aforementioned, four different ANN-based prediction models were design. Namely: M-5, M-10, M-15, and M-20. The scope of the research is limited to the twenty most recurrent diagnose, whose corresponding codes and number of related records are presented in Table 6.

Table 6. Prediction model design

ID	Diagnosis code	Number of records	M-5	M-10	M-15	M-20
C-1	R51X	2,510	✓	✓	✓	✓
C-2	A099A	1,693	✓	✓	✓	✓
C-3	A099B	1,450	✓	✓	✓	✓
C-4	U072	1,426	✓	✓	✓	✓
C-5	A09X	1,344	✓	✓	✓	✓
C-6	S525	966		✓	✓	✓
C-7	S934	926		✓	✓	✓
C-8	T384A	598		✓	✓	✓
C-9	O469	596		✓	✓	✓
C-10	A099C	569		✓	✓	✓
C-11	S525	489			✓	✓
C-12	K359	458			✓	✓
C-13	O069	449			✓	✓
C-14	R51X	422			✓	✓
C-15	ZZ01	408			✓	✓
C-16	A099D	385				✓
C-17	T384B	380				✓
C-18	O021	332				✓
C-19	T384C	320				✓
C-20	S934	319				✓

The resulting prediction models were compared by means of the percentage of correct predictions with the validation dataset. Additionally, curves Precision-Recall and the area under the curve ROC (ROC AUC) were considered during the comparison too. Davis and Goadrich provided an interesting analysis about the curves (Davis and Goadrich, 2006).

When dealing with classification problems it is important to consider the class balance. In the case of heavily imbalanced datasets, some authors recommend the inclusion of additional performance metrics such as Precision-Recall curves along with ROC AUC. Saito and Rehmsmeier explored this fact with great detail (Saito and Rehmsmeier, 2016).

3.3 Construction

By means of using six attributes, the goal was building up models capable of predicting the diagnosis of a patient received at ED. Consequently, the target class to be predicted is the attribute *Diagnosis*. Particularly, in the case of M-5 the target class can take the values of the five most common diagnose.

To overcome the limited number of attributes, six in total, the attribute *Reason for referral* was transformed from a string to a vector of words. Generating a large number of new attributes which provide additional pieces of information.

All models were developed using the well-known data processing software WEKA 3.9.6. An excellent source of information about this book about this software and functionalities was is the book by Witten et al. (Witten et al., 2017).

Initially, all models were trained and tested applying a cross-validation scheme of k=10 folds. (Table 7)

Table 7. Predictions with training and test dataset and cross-validation k=10

				Weighted average		
Model	**Codes**	**Instances**	**Correct predictions**	**Precision**	**Recall**	**ROC AUC**
M-5	5	6,723	80.63 %	0.813	0.806	0.938
M-10	10	9,678	63.05 %	0.616	0.631	0.901
M-15	15	11,404	56.02 %	0.534	0.560	0.896
M-20	20	12,840	52.34 %	0.494	0.523	0.896

Although cross-validating helps reduce the risk of overfitting, the effect of the fold partitioning remains. Powers and Atyabi provided a good explanation about this fact (Powers and Atyabi, 2012). Replicating experiments might mitigate this issue by means of using different folds in each iteration. For the purposes of this research 10 replications are run, which means that each model is trained and tested 100 times. The results of such strategy show that the standard deviation of the replications fluctuates between 1.88% and 2.22% (Table 8).

Table 8. Predictions with training and test dataset, cross-validation k=10, and 10 replications

Model	Average correct predictions	Standard deviation (10 replications)
M-5	80.40 %	2.22
M-10	63.54 %	1.78
M-15	55.83 %	1.58
M-20	52.58 %	1.88

3.4 Validation
The validation of the trained and tested ANN-based models is carried out with a set of unknown data: the validation dataset held out during the phase of analysis. This dataset contains unseen instances which are equivalent to 20% of the data preprocessed for each model.

4. Data Collection
Validation results show consistency between the results obtained with the training and test data, and the result with the validation unknown data (Table 9). Even though, prediction percentages are rather low in the cases of models M-15 and M-20, when the number of classes (or diagnosis codes) is limited to five, the results are promising.

Table 9. Correct prediction rates with unknown validation data

				Weighted average		
Model	**Codes**	**Instances**	**Correct predictions**	**Precision**	**Recall**	**ROC AUC**
M-5	5	1,700	83.56 %	0.845	0.836	0.959
M-10	10	2,400	66.58 %	0.664	0.665	0.923
M-15	15	2,900	58.86 %	0.561	0.589	0.915
M-20	20	3,200	54.78 %	0.536	0.548	0.917

5. Results and Discussion
As the number of possible values of the target class (diagnose) increases, the rate of correct prediction decreases. In spite of increasing the size of the dataset, the inclusion of more class values (diagnose) make the prediction more difficult (Table 10).

Table 10. Prediction model comparison

		Train and test dada		Validation data	
Model	**Codes**	**Instances**	**Correct predictions**	**Instances**	**Correct predictions**
M-5	5	6,723	80.63 %	1,700	83.56 %
M-10	10	9,678	63.05 %	2,400	66.58 %
M-15	15	11,404	56.02 %	2,900	58.86 %
M-20	20	12,840	52.34 %	3,200	54.78 %

5.1 Numerical Results

Confusion matrices are useful to summarize the prediction results in tables. The diagonal contains the number of instances correctly classified. The other cells present incorrect classifications (Table 11).

Table 11. Confusion matrix

	Class 1	Class 2
Class 1	Instance of class 1 classified as class 1	Instance of class 1 classified as class 2
Class 2	Instance of class 2 classified as class 1	Instance of class 2 classified as class 2

A common technique to improve the correct predictions rate is the penalization of misclassification. But it was not included in this investigation. The corresponding confusion matrices of M-5 in Table 12, and Table 13.

Table 12. Confusion matrix of M-5. Training and test dataset with cross-validation (6,273 instances)

Diagnosis Codes	C-1	C-2	C-3	C-4	C-5
C-1	**645**	746	1	1	0
C-2	210	**1852**	6	12	4
C-3	0	47	**1012**	37	57
C-4	2	32	70	**1020**	12
C-5	0	4	46	9	**892**

Table 13. Confusion matrix of M-5. Validation dataset (1,700 instances)

Diagnosis Codes	C-1	C-2	C-3	C-4	C-5
C-1	**150**	150	0	0	0
C-2	35	**379**	2	2	2
C-3	0	8	**152**	15	16
C-4	2	9	13	**287**	3
C-5	0	4	11	7	**453**

It is worthwhile to emphasize that for different ED patients the same *Reason for visit* can lead to different diagnose, increasing the chances of misclassification as can be observed in Table 14 and Table 15, which correspond to M-10.

Table 14. Confusion matrix M-10. Training and test dataset with cross-validation (9,678 instances)

	C-1	C-2	C-3	C-4	C-5	C-6	C-7	C-8	C-9	C-10
C-1	**692**	679	0	14	0	3	1	1	1	0
C-2	347	**1654**	10	9	11	9	1	20	25	3
C-3	2	34	**916**	0	36	75	4	30	5	51
C-4	270	486	1	**13**	0	0	0	2	3	0
C-5	2	18	29	0	**892**	34	125	28	4	2
C-6	6	137	163	0	18	**358**	6	39	24	20

C-7	0	3	13	0	257	14	**130**	26	6	3
C-8	1	61	21	2	23	44	14	**269**	17	30
C-9	18	124	2	1	1	15	3	18	**301**	0
C-10	0	3	33	1	5	15	0	14	0	**877**

Table 15. Confusion matrix M-10. Validation dataset (2,400 instances)

	C-1	**C-2**	**C-3**	**C-4**	**C-5**	**C-6**	**C-7**	**C-8**	**C-9**	**C-10**
C-1	**146**	152	0	4	0	0	0	0	0	0
C-2	30	**362**	2	3	0	7	2	3	10	2
C-3	0	4	**133**	0	5	28	0	9	1	11
C-4	51	97	0	**3**	0	0	0	0	0	0
C-5	1	8	2	0	**237**	16	39	11	0	2
C-6	0	45	36	0	4	**85**	6	12	2	5
C-7	0	1	0	0	48	7	**54**	5	1	1
C-8	0	20	5	0	2	13	1	**63**	2	10
C-9	1	23	0	0	0	2	0	5	**81**	1
C-10	0	3	15	0	2	13	0	11	0	**434**

6. Conclusion

Data collected during 2020 allowed researchers to have a valuable insight of daily operations at an ED. The application of adequate tools and techniques helped obtain valuable information than can be used support the decision making and to improve the management of the existing resources.

Data preprocessing and the initial analysis revealed that a significant number of the ED records were associated to a rather small number of all possible diagnose. In fact, during 2020, the five most recurrent diagnose covered almost 14% of all ED records.

ANN-based prediction models offer and interesting alternative to support management tasks such as planning and scheduling. Although counting with a rather large dataset, the models proposed in this investigation intended to minimize the use of data and processing time. Therefore, only a fraction of ED database was used. In total, only six attributes were considered to predict patient diagnose. The transformation of the attribute *Reason for visit* from a string to a vector of words supplied extra pieces of information by means of creating additional attributes that helped construct more complex models.

When working with machine learning algorithms, instead of simply holding out part of the data it is advisable to apply a cross-validation scheme to minimize the influence of the dataset split up. Averaging several results is better than having only one result. Furthermore, running replications helped reduce the bias caused by the fold partitioning. Besides that, having a set of unknown data only for validation helped confirming whether the models could generalize properly or not.

Experimental work confirmed that those models with a smaller set of target class values produced higher correct prediction rates. For example, with five class values the correct prediction rate was over 80%. In contrast, with 20 class values the rate was barely over 50%.

Another interesting fact is the consistency in the experimental results obtained when applying a cross-validation scheme, a validation set with unknown data, and a battery of replications. In particular, the predictions made with unknown data proved that at least the model with 5 class values was able to generalize well.

Finally, correct prediction rates and confusion matrices suggested that the ANN-based models with 5 and 10 class values can predict ED patient diagnose with reasonably good results.

References

Berchet, C., Emergency Care Services: Trends, Drivers and Interventions to Manage the Demand, *OECD Health Working Papers* 83, OECD Publishing, Paris, 2015.

Curtis C., Chang L, Thomas J., Bollerman, O., Pianykh, S., Machine learning for predicting patient wait times and appointment delays, *Journal of the American College of Radiology*, vol. 15, no. 9, pp. 1310-1316, 2018.

Davis, J. and Goadrich, M., The Relationship Between Precision-Recall and ROC Curves, *Proceedings of the 23rd International Conference on Machine Learning*, 2006.

Divina, F., Gilson, A., Gómez-Vela, F., Garcia, M., and Torres, J., Stacking Ensemble Learning for Short-Term Electricity Consumption Forecasting, *Energies 2018*, vol. 11, no. 4, p. 949, 2018.

Kuo Y-H., Chan, N., Leung, J., Meng H., So, A., Tsoi, K., and Graham, C., An integrated approach of machine learning and systems thinking for waiting time prediction in an emergency department, *International Journal of Medical Informatics,* vol. 139, pp. 104-143, 2020.

Miles, J., Turner, J., Jacques, R., Williams, J., and Mason, S., Using machine-learning risk prediction models to triage the acuity of undifferentiated patients entering the emergency care system: a systematic review, *Diagnostic and Prognostic Research*, vol. 4, no. 16, 2020.

Mora, J., Modelos predictivos en salud basados en aprendizaje de máquina (machine learning), *Revista Médica Clínica Las Condes*, vol. 33, no. 6, pp. 583-590, 2022.

Morano, P., and Tajani, F. (2013). Bare ownership evaluation. Hedonic price model vs artificial neural network. International Journal of Business Intelligence and Data Mining. 8(4): 340-360.

Pak, A., Gannon, B., and Staib, A., Predicting waiting time to treatment for emergency department patients, International *Journal of Medical Informatics,* vol. 145, no. 1, 2021.

Pedrero, V., Reynaldos-Grandón, K., Ureta-Achurra, J., and Cortez-Pinto, E., Generalidades del machine learning y su aplicación en la gestión sanitaria en servicios de urgencia, *Revista médica de Chil*e, vol. 149, no. 2, pp. 248-254, 2021.

Powers, D., Atyabi, A., The Problem of Cross-Validation: Averaging and Bias, Repetition and Significance. *2012 Spring World Congress on Engineering and Technology*, SCET 2012 - Proceedings. 1-5, 2012.

Saito, T., Rehmsmeier, M., The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets, *PLoS ONE 10(3): e0118432*, 2015.

Witten, I., Frank, E., Hall, M., and Pal, C., *Data Mining: Practical Machine Learning Tools and Techniques*, 4th Edition, Morgan Kaufmann, Cambridge, 2017.

Salway, R., Valenzuela, R., Shoenberger, J., and Mallon, W., Congestión en el servicio de urgencia: respuestas basadas en evidencias a preguntas frecuentes, *Revista Médica Clínica Las Condes*, vol. 28, no. 2, pp. 220-227, 2017.

Santelices, E. and Santelices, J., Descripción y análisis del sistema de red de urgencia (RDU) en Chile, recomendaciones desde una mirada sistémica, *Revista Médica Clínica Las Condes*, vol. 28, no. 2, pp. 195-205, 2017.

Tang, F., Xiao, C., Wang, F., and Zhou, J., Predictive modeling in urgent care: a comparative study of machine learning approaches, *JAMIA Open*, vol. 1, no. 1, pp. 87–98, 2018.

Vollmer, L., Decruyenaere, J., Nijsten, M., Glasner, C., and Sinha, B., Machine learning in infection management using routine electronic health records: tools, techniques, and reporting of future technologies, *Journal of Clinical Microbiology and Infection*, vol. 26, pp. 1291-1299, 2020.

Biographies

Carlos Hernández is an industrial engineer, consultant, and university professor. He earned Master of Sciences in Engineering and Doctor of Engineering from Technische Universität Braunschweig, Brunswick, Germany. He is the author of several scientific and engineering articles. He has taught lectures in Discrete Event Simulation, Engineering Economics, Corporate Finances, Data Mining and Machine Learning for engineering students. He has developed a professional career working for large multinational companies (PricewaterhouseCoopers, BHP Billiton, and Merck Sharp & Dohme). He also worked as a scientific researcher in the Institut für Produktionsmesstechnick at TU Braunschweig, Germany. His research interests include manufacturing process simulation, supply chain design and simulation, and machine learning for finances. He is a member of IEOM.

Jaime Castillo is an industrial engineer, consultant, and university professor. He earned Licentiate Degree in Forest Engineering from Universidad de Concepción, Concepción, Chile, and Master in Industrial Engineering from Universidad del Desarrollo, Santiago, Chile. He has taught lectures in Project Planning & Management, Project

Evaluation, Decision Theory, and Process Simulation for engineering students. He has extensive experience working in sustainable development projects for the local forest industry. During her academic tenure he has been appointed in different management positions and has mentored over a fifty students. His research interests include project management, logistic risk assessment, and decision theory.