Using Feature Predictive Power as a Guide for the Clinical Decision Making Process

Theophano Mitsa, Ke Lin
Department of Electrical & Computer Engineering
University of Massachusetts-Dartmouth, Dartmouth, MA 02747

ABSTRACT: *Given the hectic nature of patient care environments, it is important for the clinician to reach the correct patient diagnosis as soon as possible. One of the challenges that clinicians face is to identify which patient data are salient to the identification of the correct patient status. In this paper, we propose an algorithm that was developed for incorporation in a decision-support tool for the novice nurse. The algorithm guides the patient data collection process using an iterative two-step process: (a) at each stage of the clinical decision process and based on the current partial patient information, it identifies the most probable diagnoses (b) it prompts the nurse to measure the clinical indicators with the highest predictive power of the aforementioned diagnoses. Results are presented as to the effectiveness of two predictive power measures, mutual information and the divergence measure, in guiding the data collection process.*

KEYWORDS: predictive, saliency, clinical, decision.

1. INTRODUCTION

With the severe nursing shortages of the present health care environment as the motivating factor, the Nursing Computer Decision Support system (N-CODES) was developed at UMASS-Dartmouth as a PDA-based decision support tool for novice nurses [For04]. The system was developed to aid the novice nurse (< 3yr. experience) in clinical decision making. N-CODES does not follow a cookbook approach, instead each step of its inference engine is based on the patient input data and the user action at the previous step.

During the patient evaluation process, the N-CODES PDA device is used to collect relevant patient measurements (e.g. temperature, pain level, oxygen saturation, lung sounds, etc.), which are used to perform an appropriate patient intervention. In the inference engine of N-CODES, nursing knowledge is represented as a set of data rules (for example, if cough frequency paroxymal then do respiratory intervention) and associated meta-rules. Each data rule has an evidence-based ranking of strong, sufficient, or marginal. The nursing team has created twelve practice maps, containing the data and meta-rules for a number of clinical conditions [For04]. Each map consists of a collection of data rules and paths, representing the nursing care process.

The ultimate goal of N-CODES is to aid the clinician in his/her clinical decision making, which is a complex task, that utilizes patient informational inputs and the clinician's experience, in order to identify and manage a patient's needs. Clinical decision making models fall into two major categories; analytic models and information processing models (IP). Analytic models view clinical decision making as a process of reaching a decision by pre-specification of decision alternatives, while information processing models view it a hypothetic-deductive method characterized by (a) cue recognition, (b) hypothesis generation, (c) cue interpretation, and (d) hypothesis evaluation [Els78], [Tan87], [Wes86]. An important underlying component of both models is the collection of appropriate patient data either for the formulation of a decision alternative, in the case of analytic models, or the generation of a hypothesis, in the case of IP. A question that arises here is how we define "appropriate patient data". These are data that are salient to the identification of the correct patient diagnosis and the confirmation or ruling out of high patient-risk diagnoses. In addition, given the hectic nature of clinical care environments, these salient data must be collected as early as possible in the patient assessment process, in order to correctly guide the clinical decision making process.

In this paper, we describe a novel algorithm for identifying salient data during the clinical decision making process. At each stage of this process, the data saliency computation is based on the computation of their power in predicting the diagnoses that are the most probable at that particular stage. Two different measures are investigated for predictive power computation; mutual information and the divergence measure. The paper is organized as follows: Section 2 provides a background on predictive power computation. Section 3 describes our algorithm, while section 4 provides experimentation results and a discussion. A conclusion and description of future work are provided in section 5.

2. BACKGROUND

The computation of feature predictive power is central to our algorithm's ability in capturing data saliency regarding patient status classification. We investigated two measures for the computation of feature predictive power; mutual information and the divergence measure. Mutual information is a measure of the dependence of two random variables. Given two random variables X, Y a common definition of mutual information, based on the Kullback-Leibler measure [Mae97], is:

$$MI(X,Y) = \sum_{x,y} p_{XY} \log\left(\frac{p_{XY}(x,y)}{p_X(x)\, p_Y(y)} \right) \qquad (1)$$

where $p_x(x)$, $p_y(y)$ are the marginal probability distributions and $p_{xy}(x,y)$ is the joint probability distribution. Mutual information has been used in various applications, such as image registration [Mae97] and text categorization [Han01].

Another metric of feature predictive power we investigated is the divergence measure, which assumes a Gaussian distribution and measures how different are the values of a feature in two classes. It was used in [Upp98] to find which features are the most useful in differentiating normal lung tissue from abnormal tissue in CT lung images. The divergence measure of a feature *F* for two different classes can be computed as:

$$J(C_1, C_2, F) = (\frac{(\sigma_1 - \sigma_2)^2 + (\sigma_1 + \sigma_2) * (\mu_1 - \mu_2)^2}{2 * \sigma_1 \sigma_2} \qquad (2)$$

where σ_1 and σ_2 are the variances of this feature in classes C1 and C2 and μ_1 and μ_2 are the means of this feature in these two classes. The divergence measure of Feature F for more than two classes equals to the sum of the pair divergence measures.

$$J(C_1, C_2, C_3, F) = J(C_1, C_2, F) + J(C_1, C_3, F) + J(C_2, C_3, F) \qquad (3)$$

3. ALGORITHM DESCRIPTION

The goal of our algorithm is to enable the clinician, using the N-CODES decision support tool, to correctly reach the correct patient status classification as soon as possible. The algorithm will use the following iterative process:

1. **Data collection with no prompting.** In the beginning of the patient clinical evaluation process, the clinician measures clinical indicators without any prompting from the tool.
2. **Display the most probable diagnoses.** As soon as enough patient clinical indicators are measured, the N-CODES PDA tool will display the most probable diagnoses. This computation is based on the

Bayes aposteriori probability [Upp98]. Specifically, the probability that the patient's clinical diagnosis is c (e.g., pulmonary emphysema), given clinical indicator f (e.g. cough), can be expressed as:

$$P(c/f) = \frac{P(c)P(f/c)}{P(f)} \qquad (4)$$

It is important to note that this computation will not be performed during the clinical assessment practice. Instead, the a-posteriori probabilities of various clinical diagnoses, given various combinations of clinical indicators and various values of these clinical indicators will be computed offline and stored in the N-CODES database. For example, we will compute the a-posteriori probability of pneumonia only given cough, only given paroxymal cough, given cough and sputum production, and paroxymal cough and sputum production.

3**. Display the clinical indicators with the highest predictive power of the most probable diagnoses.** The predictive power can be computed using either the divergence measure shown in Equation (2) or the mutual information measure. Because many clinical indicators are either binary (for example skin rash present versus skin rash non-present) or of the categorical type (for example, low fever and high fever), we have to use an appropriate definition of mutual information. Given a binary clinical indicator f and a class c, the mutual information definition we use in our algorithm was adopted from [Han01] and is given by:

$$MI(f,c) = P(c,f)\log\frac{P(c,f)}{P(c)P(f)} + P(c,\bar{f})\log\frac{P(c,\bar{f})}{P(c)P(\bar{f})} \qquad (5)$$

where $P(f)$ is the probability of the presence of the feature and $P(\overline{f})$ is the probability of the absence of the feature. In the case of a categorical type feature, we first assign discrete values to the different categories and then compute the mutual information as follows:

$$MI(F,c) = average\ MI(f_i,c) \text{ where } f_i \text{ is one of values of } F \qquad (6)$$

In the case of more than two classes, $c_1, c_2, \ldots, c_n$, mutual information of feature f is computed as

$$MI(C_1, C_2, \ldots, C_n) = average\ MI(F, c_i) \qquad (7)$$

4. **Iterate.** Repeat steps 2 and 3 until one of the diagnoses has a significantly higher probability than the others, in which case it is chosen as the patient status.

4. RESULTS AND DISCUSSION

Prior to incorporating the algorithm in the N-CODES tool for clinical use, we simulated the data collection process using a clinical database. The purpose of this simulation was twofold: (a) to determine the success rate of our algorithm regarding correct case classification (b) to evaluate the classification speed of our algorithm. Here is a description of the database that we used:

Clinical database. The IMPROVE digital library [Tho97] is an intensive care database of 59 patients monitored in an interval of 24 hours. In addition for 7 patients, EEG signals have also been recorded. The aim of the IMPROVE project was to develop methods of bio-signal processing and interpretation which may help to detect disorders of oxygen delivery to vital tissues early enough for effective treatment. The four disorders included in the study are Hypovolaemia, Cardiac failure, High flow state, and content related problem. The data consist of 87 cases with 14 attributes. Our classifier used 86 cases for training data, and 1 for testing in a circular fashion; i.e., all cases became test cases while the remaining cases became the training data. The steps below describe how we simulated our algorithm on the clinical database. Specifically, for each case in the testing data:

Step 1: Randomly select two attributes

Step 2: Depending on the value of those attributes, determine the three classes c_1, c_2, c_3 which have the highest a-posteriori probability, in comparison with the other classes. Note that only two classes could have a significantly higher probability than the other classes.

Step 3: Find the attribute F that has the highest Divergence Measure $J(C_1, C_2, C_3, F)$ or Mutual Information $MI(C_1, C_2, C_3)$ of classes c_1, c_2, c_3. If only two classes have a significantly higher probability, then we compute $J(C_1, C_2, F)$ or $MI(C_1, C_2)$.

Step 4. Repeat Steps 2 and 3 until the highest class probability exceeds a certain threshold above the second highest class probability or until there are no more attributes to select.

Table 1 shows the results of applying our algorithm to the classification of cases from the clinical database. Specifically it shows the correctly classified test cases, incorrectly classified test cases, and unclassified cases. A case is unclassified at the end of the process, if no class probability is significantly higher than other class probabilities. The second column shows the results of using the mutual information as the predictive power measure and the third column shows the results of using the divergence measure. The table also shows the number of steps required to reach the end of the classification process. In addition, the table shows the result of selecting the next feature randomly instead of using the predictive power as a selection criterion.

	Mutual inf.	**Div. measure**	**Random**
Correctly classified	**76**	**72**	**66**
Incorrectly classified	**0**	**0**	**0**
Not classified	**11**	**15**	**21**
Av. # of steps	**3.632**	**3.847**	**4.788**

Table 1. Clinical database results

As can be seen from the table above, the divergence measure and mutual information produce comparable results, with the mutual information producing a slightly higher classification rate and slightly higher classification speed. Randomly selecting the next feature, results in a higher number of unclassified cases and also an increased number of steps.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we presented an algorithm for guiding the clinical data collection process, such that the clinician reaches the correct patient diagnosis in as timely manner as possible. Guiding the data collection process is particularly important for the novice nurse that does not have the experience to identify the salient data in a maze of possible clinical data to be collected. The algorithm uses a-posteriori probabilities computed from the input data to determine the most probable case classification. It then finds the feature with the highest predictive power, which becomes the next feature to be measured. We simulated our algorithm on a clinical database, and our results indicate that guiding the data collection process through the predictive power measures, divergence measure and mutual information, produced better results both in terms of classification rate and classification speed than random feature selection. Our next step is to incorporate this algorithm in the N-CODES tool and try it in a clinical environment.

6. ACKNOWLEDGEMENTS

We would like to thank Sweeta Brahme for the clinical data preparation and Professors Paul Fortier and Nancy Dluhy for their valuable feedback.

REFERENCES

[Els78] Elstein, A. S., L. Schulman, S. A. Sprafka, "Medical Problem Solving: An Analysis of Clinical Reasoning", Cambridge: Harvard University Press, 1978.

[For04] Fortier, P., H. Michel, N. Dluhy, E. O'Neill, "Computerized Decision Support Aid for the Novice Nurse", IADIS International Journal of the WWW/Internet, vol. II, no.2, pp. 29-42, December 2004.

[Han01] Han, E. H., G. Karypis, and V. Kumar, "Text Categorization Using Weight Adjusted K-Nearest Neighbor Classification", Lecture Notes in Computer Science, vol. 2035, 2001.

[Mae97] Maes, F., A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality Image Registration by Maximization of Mutual Information", IEEE Transactions on Medical Imaging, vol. 16, no. 2, pp.187-198, April 1997.

[Tan87] Tanner, C. A., K. Padrick, U. Westfall and D. Putzier, "Diagnostic Reasoning Strategies of Nurses and Nursing Students", Nursing Research, 36, pp. 358-363, 1987.

[Tho97] Thomsen, C.E. et al., "Collecting EEG Signals in the IMPROVE Data Library", IEEE Engineering in Medicine and Biology Magazine, vol. 16, no. 6, pp. 33-40, 1997.

[Upp98] Uppaluri R., T. Mitsa, E. A. Hoffman, G. McLennan, and M. Sonka, "Method and Apparatus for Analyzing CT images to Determine the Presence of Pulmonary Tissue Pathology", U. S. patent, # 6,466,687, 10/15/02.

[Wes86] Westfall, U., C. Tanner, D. Putzier, and K. Padrick, "Activating Clinical Inferences: A Component of Diagnostic Reasoning in Nursing", Research in Nursing and Health, 9, pp. 269-277, 1986.