A DATA-DRIVEN APPROACH FOR MATCHING CLINICAL EXPERTISE TO INDIVIDUAL CASES

Onur Atan, Cem Tekin and Mihaela van der Schaar

UCLA Electrical Engineering

William Hsu

UCLA Radiology Sciences

ABSTRACT

Hospitals are increasingly utilizing business intelligence and analytics tools to mine electronic health data to uncover inefficiencies in care delivery (e.g., slow turnaround times, high readmission rates). Given that the expertise and experience of healthcare providers may vary significantly, an area of potential improvement is optimizing the way patient cases are recommended to clinical experts (e.g., the pathologist who is most adept at diagnosing a rare cancer). In this paper, we propose an expert selection system that automatically matches a given patient case to the best available expert considering both the available contextual information about a patient (e.g., demographics, medical history, signs and symptoms, past interventions) and the congestion of the expert. We prove that as the number of patients grows, the proposed algorithm will discover the best expert to select for patients with a specific context. Moreover, the algorithm also provides confidence bounds on the diagnostic accuracy of the expert it selects. While the proposed system can be applied in many scenarios, we demonstrate its performance in the context of assigning mammography exams to individual radiologists for interpretation. We show that our proposed system can improve current clinical practice by improving overall sensitivity and specificity of screening exams compared to random assignment.Finally, since each expert can only take a certain number of diagnosis decisions on a daily basis, we show how our system can take the experts' workload into account as well as the expertise when deciding how to select experts.

1. INTRODUCTION

A pressing challenge in today's healthcare environment is matching the increased capacity to generate clinical data with a comparable ability to analyze and derive insights from this data to improve the efficiency and accuracy of care delivery. The utilization of healthcare informatics tools and decision support systems to mine patterns and quality metrics from structured clinical data is vital. Current standard of practice is largely guided by evidence provided by randomized clinical trials and published as systematic reviews or clinical practice guidelines.A drawback of current evidence sources is that they do not take into account the diversity in terms of patients and their health states, level of expertise exhibited by healthcare professionals, and the tests and equipment available at different institutions. [1] For instance, studies have shown that less experienced radiologists have higher false-positive rates than experienced radiologists when detecting breast cancer from mammography images, and the detection accuracy over different healthcare providers ranges from 2.6% to 15.9% [2]

One approach towards compensating for these differences and providing uniform care to patients is the implementation of data-driven *Clinical Decision Support Systems* (CDSSs). These systems have been demonstrated to improve a variety of detection and diagnostic tasks in areas such as lung cancer [3], breast cancer [4, 5] and diabetes [6]. Although recent surveys show that CDSSs have improved the accuracy of clinical decision making process for some applications [7], their effectiveness may vary significantly depending on institution and patient population [8]. The success of a CDSS or a human expert is based on many factors, including the complexity of the diagnostic problem, the training of the personnel using the CDSS, and the ease of use of the CDSS by healthcare professionals. One of the key challenges faced by health clinics is deciding which CDSS to use, understanding under what circumstances should a CDSS be used, optimizing the selection of individual healthcare professionals based on cases in which they achieve the highest diagnostic performance.

Complementing prior efforts to develop CDSSs for different diagnostic tasks, this paper studies the broader problem of managing the clinical workflow by dynamically assigning unseen patient cases to the appropriate expert. Here, expert collectively refers to multiple sources for obtaining a diagnosis: a clinician, a CDSS, or or a combination of them (e.g., using the CDSS as a first-pass reader then having an expert review the results). Our proposed system for discovering expertise exploits the fact that the diagnosis accuracy can be accurately characterized and predicted using a set of *relevant* contexts of the patient. While in practice, contexts may include information that is not explicitly documented in the patient record (e.g., a physician's experience, subtle features seen in a radiologic image), we initially limit our modeling to structured information in the medical record that can be utilized in the decision making process of the expert. Our methodology uses the context of the patient to estimate the level of expertise exhibited by the expert and assigns the expert by optimizing the trade off between their expertise and congestion cost. The level of the expertise is defined based on the accuracy of their diagnostic and congestion cost is defined as a function ofthe number of patient cases assigned to them. An important challenge is the dimensionality of information contained within the context of a patient, which includes vast set of features that are irrelevant to the diagnosis of the expert. Filtering out these features is crucial for both improving the learning speed of the expertise, i.e., learning accurately from a small number of patients, and reducing the time it takes to diagnose, i.e., showing only the relevant contexts to the expert, hence reducing the time the expert needs to analyze and figure out which context is relevant and which is not. In the proposed method, the relevant contexts are learned online, based on past patient data and expert recommendations, and expertise recommendations are adapted based on the relevant contexts.

In summary, this paper proposes a novel algorithm that (*i*) learns online the diagnostic accuracy of different human experts and/or CDSSs for different patient types and cases, (*ii*) assigns an expert to each specific case based on their estimated diagnostic accuracies and current congestion costs, (*iii*) discovers the contexts that are relevant to the diagnosis and filters out the irrelevant contexts, (*iv*) learns the current diagnostic performance of experts and can automatically

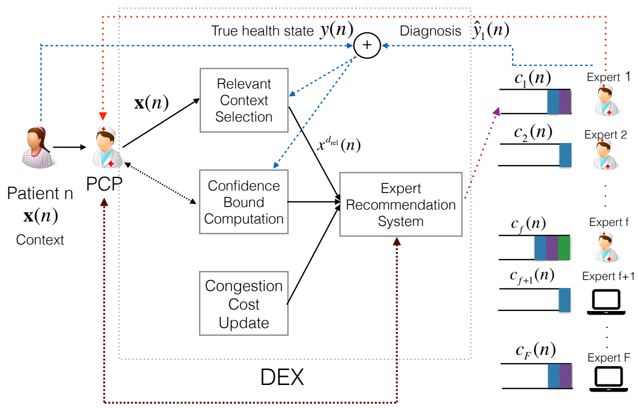

Fig. 1. Expertise Recommendation System

adjust the assignment of new cases (e.g., to compensate for fatigue) or alert the primary care physician (PCP) ordering the exam about potential issues.

For proposed algorithm, we formally (i) prove that it converges to the optimal expert recommendation, (ii) provide confidence bounds on the diagnostic accuracy of the recommended expert.

2. RELATED WORK

Most of the prior work does not consider the context of the patient, which can be easily extracted from their electronic health records. A key challenge for the works that do consider contextual information is that the contexts are high-dimensional and hence, the context-driven decisions are plagued by the curse of dimensionality. For example, [9] proposes penalizing the irrelevant contexts with an L1 norm cost. However, they assume a linear dependence between the context and labels. Our methodology is more general and does not require linearity.

Some of the existing solutions [10, 11, 12, 13, 14] propose ensemble learning techniques. These works show asymptotical convergence results to the optimal without any convergence rate. On the contrary, we provide confidence levels to the selection of our method, which is especially important for the health care systems.

Our proposed method builds on and significantly extend contextual bandits, which have been introduced and studied before in [15, 16, 17, 18]. The key differences are that our method (i) learns the relevant contexts, and decides experts to assign based on them, which results in significantly faster learning, (ii) tradeoffs diagnostic accuracy and congestion, (iii) learns the expertise levels by safe explorations, which are essential to ensure that the patient safety is not compromised.

3. PROBLEM FORMULATION

The set of all experts including human experts and CDSSs are denoted by $\mathcal{F}$. The number of experts is $F = |\mathcal{F}|$, where $|\cdot|$ is the cardinality operator. Patients arrive sequentially and are indexed according to their arrival times by $n = 1, 2, \ldots, N$. The context vector of patient n is $\boldsymbol{x}(n) = (x^1(n) \ldots x^D(n))$, where D is its dimension,[1] $x^d(n)$ is the type-d context, and $\mathcal{D} =: \{1, \ldots, D\}$ is the set of context types. The space that type-d contexts belong to is denoted by $\mathcal{X}^d$, and the (joint) context space is denoted by $\mathcal{X} = \times_{d \in \mathcal{D}} \mathcal{X}^d$. Without loss of generality we assume that the contexts are normalized, i.e., $\mathcal{X}^d = [0, 1]$ for all $d \in \mathcal{D}$. However, our algorithm will work and our results will hold for any discrete or continuous $\mathcal{X}^d$ provided that is is bounded.

For the nth patient the following events happen sequentially: (i) the context vector $\boldsymbol{x}(n)$ of the patient is extracted from an electronic health record and/or by evaluation of the PCP or a nurse ; (ii) based on the estimated diagnostic accuracies and the current congestion levels of the experts, the algorithm recommends one of the experts $a(n) \in \mathcal{F}$ to the PCP, and outputs a confidence bound on the diagnostic accuracy of the recommended expert for $\boldsymbol{x}(n)$, (iii) upon evaluating the case, the recommended expert recommends a diagnosis $\hat{y}_{a(n)}(n) \in \mathcal{Y}$ for the patient, where $\mathcal{Y}$ is the set of possible diagnosis recommendations (in the application of breast cancer, this set includes breast tumor being malignant or benign); (iv) after some delay, the true health state of patient $y(n) \in \mathcal{Y}$ is revealed,[2] ;(v) the algorithm adapts its expert recommendation strategy. See figure 1 for pictorial description of our system.

For each patient n with context vector $\boldsymbol{x}(n)$, let $\pi_f^d(x^d(n))$ be the *marginal diagnostic accuracy* (or simply, accuracy) of expert $f \in \mathcal{F}$, where $f(x^d(n))$ is the diagnosis recommendation of expert f when only type-d context information of the patient is shown to the expert. The diagnostic rule used by expert f, i.e., $f(\cdot)$, can be either deterministic or random, and is not required to be known by the system.

An expert will make similar diagnosis decisions for patients with similar contexts [19]. This is formalized as follows.

Definition 1. ***Similarity Property.*** *For each $f \in \mathcal{F}$ and $d \in \mathcal{D}$, there exists constants $\alpha > 0$ and $L > 0$, such that for all $x, x' \in \mathcal{X}^d$, we have $|\pi_f^d(x) - \pi_f^d(x') \leq L|x - x'|^\alpha$.*

For our theoretical analysis to hold, our algorithm requires a lower-bound on α as an input parameter, while knowledge of L is not required. Such a lower bound can be estimated from past empirical data as well as from opinions of experts in the field. In our experimental results in Section 5, we examine the affect of different values of α (as an input parameter) on the performance of our algorithm. The accuracies $\pi_f^d(x^d)$ for any $f \in \mathcal{F}$, any $x \in \mathcal{X}_d$ and $d \in \mathcal{D}$, are unknown a priori.

Making a diagnosis for a new patient requires a certain amount of time devoted by the expert. Hence, an expert gets congested when too many new patients are recommended, which results in delay in making a diagnosis. We model this as the congestion cost $c_f(n)$, which is given as $c_f(n) = \min\{c_f(n-1) + u_{\boldsymbol{x}(n)} I(\alpha(n) = f) - p_f \delta(n-1, n))^+, 1\}$, where $p_f \in [0, 1]$ is the amount of work that expert f can do in one unit of time, $\delta(n-1, n)$ is the time between the arrival of patient $n-1$ and n, and $u_{\boldsymbol{x},f} \in [0, 1]$ is the amount of work to diagnose the patient with context $\boldsymbol{x}$ and $c_f(0) \in [0, 1]$ for all $f \in \mathcal{F}$. We assume that $c_f(n)$ is known (as u_x and p_f can be calculated from the historical patient data).

For expert f and patient n, the *diagnosis reward* is defined as $r_f(n) := \mathrm{I}(\hat{y}_f(n) = y(n))$, which is 1 if the diagnosis is correct and 0 otherwise. This reward is observed only after the true health state $y(n)$ is revealed.

The goal of our systemis to maximize the total *net reward* for the patients that have arrived so far, where the *net reward* of a patient n is defined as $r_f(n) - \lambda c_f(n)$, where the constant $\lambda > 0$ set by the clinic captures the tradeoff between accuracy and congestion (delay). This corresponds to minimizing the regret with respect to the benchmark solution (which has the same tradeoff factor). Our benchmark is the solution which selects the expert in $\mathcal{F}$

[1]Each dimension represents a different type of context. For example, first dimension may represent age, second dimension may represent weight, third dimension may represent gender, etc.

[2]Our algorithm will also work when the true health state of some patients is never recovered by simply disregarding the history related to that patients.

with the highest net reward given the context vector $\boldsymbol{x}(n)$ for patient n. Specifically, the solution we compare against is given by $f^*_n(\boldsymbol{x}) := \arg\max_{f \in \mathcal{F}} \left(\max_{x^d \in \boldsymbol{x}} \left(\pi^d_f(x^d) - \lambda c_f(n) \right) \right), \quad \forall \boldsymbol{x} \in \mathcal{X}$ and $\forall n = 1, 2, \dots N$. Since $f^*_n(\boldsymbol{x})$ calculated based only on the marginal accuracies of the experts for the patient n and their congestion level, we call $f^*_n(\boldsymbol{x})$ as the *the marginal best available expert* given the patient's context $\boldsymbol{x}(n)$. Different from prior work [15, 16], in which the recommendation depends on the context of the patient, in our benchmark, the recommendation also depends on the patient index n, since different patients experience different congestion levels. Hence, the benchmark may be different for the patients with the same contexts. This benchmark is the optimal policy among all the policies that shows the experts only the relevant context of the patients, instead of the entire context vector.

We define the regret to be the loss incurred due to the unknown expertise. Regret of a learning algorithm which recommends an expert $a(n) \in \mathcal{F}$ for patient n is defined as

$$\begin{aligned} \mathrm{Reg}(N) := &\sum_{n=1}^{N} \pi_{f^*_n(\boldsymbol{x}(n))} - \lambda c_{f^*_n(\boldsymbol{x}(n))}(n) \\ &- \mathrm{E}\left[\sum_{n=1}^{N} r_{a(n)}(n) - \lambda c_{a(n)}(n) \right]. \end{aligned}$$

Regret gives the convergence rate of the total expected reward of the learning algorithm to the value of the benchmark solution $f^*_n(\boldsymbol{x})$, $\boldsymbol{x} \in \mathcal{X}$.

4. DISCOVER THE EXPERT (DEX)

The basic idea behind our *Discover the Expert* (DEX) algorithm is to learn the accuracies of different experts while requesting diagnosis recommendations from them in an efficient way. There are two operation phases of DEX: safe exploration and exploitation. In an exploitation phase, DEX is very confident about its expert selection decision and recommends the expert with the highest estimated net reward. In the safe exploration phase, DEX tries an expert whose net reward is not known accurately, i.e., an expert with a high uncertainty in the estimated diagnostic accuracy. Hence, in a safe exploration phase, DEX can alert the PCP when the overall diagnostic performance of available experts is reduced or congested for the given patient and context, in which case the PCP may decide to consult another expert or utilize a combination of CDSS and expert In both phases, the estimated accuracy of the selected expert is updated after the true health state of the diagnosed patient is revealed.

DEX adaptively divides the context space into finer and finer regions as more patients arrive such that the regions of the context space with large number of arrivals are explored more accurately than regions of the context space with small number of arrivals, and then only uses the observations in those sets when estimating the diagnostic accuracies of the experts.

DEX keeps separate counters, estimates and intervals of context for each type- d context to learn the relevant context type.

Intervals of Contexts Let $\mathcal{P}_l(\mathcal{C})$ be the uniform partition of a 1-dimensional space $\mathcal{C}$ formed by intervals of length 2^{-l+1} intervals, where l is called the level of the partition. The set of all intervals which can be elements of a partition of $\mathcal{X}^d$ is $\mathcal{P}^d = \cup_{l=1}^{\infty} \mathcal{P}_l(\mathcal{X}^d)$.

Active Context Subspace. For the patient n and type of context-d, DEX keeps an *active* partition of $\mathcal{X}^d$ denoted by $\mathcal{A}^d_n$, such that $\cup_{C \in \mathcal{A}^d_n} C = \mathcal{X}^d$. Let $C^d(n)$ denote the interval in $\mathcal{A}^d_n$ that the type-d context $x^d(n)$ lies in, and $\boldsymbol{C}(n) = (C^1(n) \dots C^D(n))$ be the set of intervals in which context $\boldsymbol{x}(n)$ lies in. The expert recommendation decision for patient n is done only based on the past information from patients whose contexts lie in $\boldsymbol{C}(n)$.

Activation, Partitioning and Deactivation. Once an interval $C^d \in \mathcal{P}^d$ is activated (becomes part of $\mathcal{A}^d_n$, we maintain a counter T^d_C that records number of patients with contexts arriving to C^d beginning from its activation. A level l interval C^d stays active until the first patient n such that $T^d_C > A2^{pl}$ (where $p > 0$ and $A > 0$ are the algorithm input parameters), which is to ensure that the number of active hypercubes are logarithmically many in the number of patients. Then, it is further partitioned into 2 intervals, which are denoted by $\mathcal{P}_{l+1}(C^d)$. The intervals on the set $\mathcal{P}_{l+1}(C^d)$ are activated while the interval C^d is deactivated. The set of active intervals for the type-d contexts for the patient n becomes $\mathcal{A}^d_n = (\mathcal{A}^d_n \setminus C^d) \cup \mathcal{P}_{l+1}(C^d)$.

Let $T^d_{C^d}(n)$ be the number of patients that arrived before patient n, whose type-d contexts are contained in the interval C^d (starting from its activation). For $f \in \mathcal{F}$, let $T^d_{C^d,f}(n)$ be the number patients before patient n that are recommended to expert f, whose type-d contexts are contained in C^d (starting from its activation).

When patient n arrives, DEX first identifies $\boldsymbol{C}(n)$. Then, it determines whether to explore or exploit. Let $\mathcal{S}^d_C(n) := \{f \in \mathcal{F} : T^d_{C^d,f}(n) \le H(n)\}$ where $H(n)$ is a control function for the patient n, which determines whether the experts are explored sufficiently many times among the similar patients. Let $S_{\boldsymbol{C}}(n) = \cup_{d \in \mathcal{D}} S^d_C(n)$. If there exists $d \in \mathcal{D}$ such that $\mathcal{S}_{\boldsymbol{C}}(n) \neq \emptyset$, then DEX recommends the least congested expert in this set, i.e $a(n) \in \min_{f \in S^d_{C^d}(n)} c_f(n)$. If $S_{\boldsymbol{C}}(n) = \emptyset$, DEX first discovers the relevant context $d_{\text{rel}} \in \arg\max_{d \in \mathcal{D}} \bar{r}^d_{C^d,f}(n) - \lambda c_f(n)$, then recommends the expert based on the relevant context type d_{rel}, $a(n) \in \arg\max_{f \in \mathcal{F}} \bar{r}^{d_{\text{rel}}}_{C^{d_{\text{rel}}},f}(n) - \lambda c_f(n)$, where $\bar{r}^d_{C^d,f}(n)$ is the sample mean diagnostic accuracy of the expert f among the patients whose type-d context lies in the hypercube C^d. Define $\mathcal{E}^d_{C^d,f}$ as the set of rewards that are collected through past patients whose context lies in C^d and are recommended to expert f. Then, marginal sample mean diagnostic action rewards can be computed as $\bar{r}^d_{C^d,f} := (\sum_{r \in \mathcal{E}^d_{C^d,f}} r)/(|\mathcal{E}^d_{C^d,f}|)$.

Before stating our formal results, we begin with a few definitions and introduce a few notations. Let $\overline{\mu}^d_{C^d,f}(n) := \sup_{x^d \in C^d} \pi^d_f(\boldsymbol{x}) - \lambda c_f(n)$ and $\underline{\mu}^d_{C^d,f} := \inf_{x^d \in C^d} \pi^d_f(\boldsymbol{x}) - \lambda c_f(n)$ and $\overline{\mu}_{\boldsymbol{C},f}(n) = \max_{x^d \in \boldsymbol{x}} \overline{\mu}^d_{C^d,f}(n)$ and $\underline{\mu}_{\boldsymbol{C},f}(n) = \max_{x^d \in \boldsymbol{x}} \underline{\mu}^d_{C^d,f}(n)$. For the patient whose context lies in the hypercube C, the suboptimal experts are given by

$$\mathcal{L}_{\boldsymbol{C}}(n) := \left\{ f \in \mathcal{F} : \underline{\mu}_{\boldsymbol{C},f} - \overline{\mu}_{\boldsymbol{C},f} > Bn^{\theta} \right\},$$

where $B > 0$ and $\theta < 0$ are parameters that are used only in the analysis of the regret and do not need to be known by the clinic. When the context vector of the patient is in C, any expert that is not in $\mathcal{L}_{\boldsymbol{C}}$(n) is a near-optimal expert.

Theorem 1. *The regret of DEX is* $O\left(DFN^{f(\alpha)} \log N\right)$ *where* $f(\alpha) = \frac{2+\alpha+\sqrt{9\alpha^2+8\alpha}}{2+3\alpha+\sqrt{9\alpha^2+8\alpha}}$.

Theorem 1 gives a bound on the long-term performance of DEX. It shows that the regret bound is sub linear in the number of patient arrivals which guarantees that DEX converges to the *marginal best available expert*. This implies that the convergence rate to the *marginal best available expert* is $O(N^{1-f(\alpha)})$. Notice that this does

Initialization: $\mathcal{A}_1^d = \mathcal{P}(\mathcal{X}^d), T_{C^d}^d(1) = 0, \bar{r}_{C^d,f}^d(1) = 0$, $T_{C^d,f}^d = 0, c_f(1) = 0 \, \forall f \in \mathcal{F}, \forall C^d \in A_1^d$ and $\forall d \in \mathcal{D}$

while $n \geq 1$ **do**

  Update the congestion cost $c_f(n) \, \forall f \in \mathcal{F}$

  Find hypercube C^d that $x^d(n)$ lies in

  $T_C^d(n) + + \, \forall d \in \mathcal{D}$

  if $\exists d \in \mathcal{D}$ s.t. $\mathcal{S}_{C^d}^d(n) \neq \emptyset$ **then**

    Explore the relevance of context type $d_{\text{rel}} = d$.

    Recommend expert $a(n) \in \min_{f \in \mathcal{S}_{C^d}^d(n)} c_f(n)$.

  else

    Discover the relevant context

    $d_{\text{rel}} \in \arg\max_{d \in \mathcal{D}} \bar{r}_{C^d,f}^d(n) - \lambda c_f(n))$

    Select diagnostic action

    $a(n) \in \arg\max_{f \in \mathcal{F}} \bar{r}_{C^{d_{\text{rel}}},f}^{d_{\text{rel}}}(n) - \lambda c_f(n))$.

  end if

  Receive diagnostic recommendation $\hat{y}_{a(n)}(n)$.

  Receive true health state $y(n)$

  Compute reward $r_{a(n)}(n) = \mathrm{I}(\hat{y}_{a(n)}(n) = y(n))$.

  $$\bar{r}_{C^{d_{\text{rel}}},a(n)}^{d_{\text{rel}}} = \frac{T_{C^{d_{\text{rel}}},a(n)}^{d} \bar{r}_{C^{d_{\text{rel}}},a(n)}^{d_{\text{rel}}} + r_{a(n)}(n)}{T_{C^{d_{\text{rel}}},a(n)}^{d_{\text{rel}}} + 1}.$$

  $T_{C^{d_{\text{rel}}},a(n)}^{d_{\text{rel}}} + +$.

  if $\exists d \in \mathcal{D}$ s.t.$T_{C^d}^d(f) > A2^{pl}$ **then**

    $\mathcal{A}_n^d = (\mathcal{A}_n^d \setminus C^d) \cup \mathcal{P}_{l+1}(C^d)$

  end if

end while

Fig. 2. Pseudocode of the DEX algorithm.

not depend on the dimension of the context since only the *relevant* context is used when recommending the experts. It also shows that the regret increases linearly with number of experts. We note that the regret is the gap between the total expected reward of the optimal policy and the total expected diagnostic reward of DEX. Since the performance of optimal policy never gets worse as more experts are introduced, the benchmark will improve. Therefore, the total reward of DEX will improve even if the regret increases with F.

Corollary 1. *The average learning loss goes to zero, i.e.,* $\lim_{N \to \infty} Reg(N)/N = 0$.

Corollary 1 shows that the loss due to not knowing the expertise (compared to the optimal policy that knows the accuracies of all the experts for all contexts) goes to zero as the number of patients increase, which is a direct result of Theorem 1.

Theorem 2. *If DEX is in exploitation phase for patient n, then with probability* $\min\left\{0, 1 - \frac{2}{n^2}\right\}$*, the recommendation is made by a near optimal expert in set* $\mathcal{K} - \mathcal{L}_{\boldsymbol{C}}(n)$.

Due to space limitation, proofs of Theorem 1 and 2 is given in online appendix. [20] In a clinical setting, for interpreting the diagnosis recommendation provided by DEX, clinicians may want to know the confidence about the proposed diagnosis recommendation for the patient under consideration. Theorem 2 shows that DEX can provide the clinicians sharp confidence bounds on the diagnostic accuracy of the expert it selects. These bounds reveal the context-specific expertise level of the human experts or CDSSs, e.g. a CDSS may excel at diagnosing cancers in a specific type of patients (having dense breasts) while another may excel at another type.

5. ILLUSTRATIVE RESULTS

To illustrate the performance of our proposed system, we consider a breast cancer data set provided by UCLA radiology department de-identified set of data on 11050 patients who undergo breast screening seen at our institution's medical center.A radiologist interprets the breast image of the patients and assigns a BI-RADS score. Score '1' is negative, '2' and '3' are associated with benign, '4' is suspicious, '5' is highly probable malignancy, and '6' is known malignancy. The diagnostic decision of the medical expert is considered using this interpretation. The dataset has following information which is used as the context: patient's age, imaging modality and breast density information.

Comparison of diagnostic accuracy : Table 1 provides a comparison of DEX against a random selection strategy, which we call "worklist assignment", representing a simplification of how experts are assigned to cases today. As seen from the table, DEX outperforms the random recommendation strategy by 20% in terms of the diagnostic accuracy. We also compare DEX with a version of DEX that does not adaptively generate the context partition, which is called DEX no context (DEX-NC). It is observed that DEX also outperforms DEX-NC by 10% in terms of the diagnostic accuracy.

Table 1. Comparison of DEX with other algorithms

Methods	DEX	DEX-NC	Worklist
Diagnostic Accuracy	81.17%	73.41%	68.45%

Results on congestion cost Table 2 and 3 provide a comparison of the diagnostic accuracy and the worklist assignment of the experts under two different models (i) a model that assumes congestion cost, and (ii) a model that does not consider congestion. Table 2 shows the fraction of times the expert 1, 2, 3 (the most congested experts) and rest of them (Expert 4,5,6) are recommended. We simulate this by taking congestion cost $u_{\boldsymbol{x},f} = 0.4$,$p_f = 1$, $\lambda = 1$ for all $\boldsymbol{x} \in \mathcal{X}$ and $f \in \mathcal{F}$ and arrival times $\delta(n-1, n) \sim \exp(\beta)$. Smaller values of β imply more frequent patient arrivals.

Table 2. Allocation of the Experts under congestion cost

Allocation	No congestion	$\beta = 0.2$	$\beta = 0.1$
Expert 1	67.39%	28.95%	21.87%
Expert 2	12.80%	27.19%	20.90%
Expert 3	7.25%	20.36%	18.15%
Other Experts	12.56%	23.5%	39.08%

As seen from the tables, the performance of DEX is minimally affected and workload becomes more uniform among the experts when the congestion costs are introduced. It also shows that by setting the trade-off can be made between the congestion cost and diagnostic accuracy.

6. CONCLUSIONS

In this paper we proposed a context-adaptive medical diagnosis system that selects from a pool of human experts and CDSSs to make diagnosis recommendations. The system learns online, which context of the patient to use, and which expert to rely on when making diagnosis recommendations. We prove that the diagnostic accuracy of the proposed system converges to the accuracy of the best context-adaptive expert. The proposed mechanism can also be used for recommending and selecting experts in other settings in healthcare settings and beyond.

Table 3. Resulting performance of DEX under congestion cost

Congestion	No congestion	$\beta = 0.2$	$\beta = 0.1$
Diagnostic Accuracy	81.17%	79.90%	77.93%

7. APPENDICES

7.1. Preliminaries

We start with a simple lemma which gives an upper bound on the highest level hypercube that is active for any patient n.

Lemma 1. ***A bound on the level of active hypercubes.*** *All the active hypercubes $\mathcal{A}^d(n)$ for type-d contexts for patient n have at most a level of $(\log_2 n)/p + 1$.*

Proof. Let $l+1$ be the level of the highest level active hypercube. We must have $A\sum_{j=0}^{l} 2^{pj} < n$, otherwise the highest level active hypercube will be less than $l+1$. We have for $n/A > 1$, $A\frac{2^{p(l+1)}-1}{2^p-1} < n \Rightarrow 2^{pl} < \frac{n}{A} \Rightarrow l < \frac{\log_2 n}{p}$. ☐

Lemma 2. ***Regret of explorations in a hypercube.*** *The total regret in such a hypercube due to safe explorations up to the nth patient of clinic is bounded by $O(DFn^z \log n)$. This is achieved by setting $D(n) = n^z \log n$.*

Proof. This directly follows from the number of trainings and explorations that are required before any diagnostic action can be exploited (see definition of $S_{\boldsymbol{C}(n)}(n)$). If the diagnosis recommendation at any exploration phase is incorrect or an expert with high congestion cost is recommended, clinic loses at most 2 from the highest realized reward it could get for that patient, due to the fact an incorrect diagnosis recommendation will result in one unit of loss and the congestion cost can at most be one. ☐

Lemma 3. ***Regret due to suboptimal expert selections in a hypercube.*** *The regret due to suboptimal expert selection in an hypercube is bounded by $\frac{2D}{n^2}$ in the exploitation phase when $2H_n \leq (B-2L)2^{-l_{\max}(\boldsymbol{C})\alpha}$ holds.*

Proof. Let Ω denote the space of all possible outcomes, and w be a sample path. The event that the DEX exploits when $\boldsymbol{x}(n) \in \boldsymbol{C}$ is given by $\mathcal{W}_{\boldsymbol{C}}(n) := \{w : S_{\boldsymbol{C}}(n) = \emptyset, \boldsymbol{x}(n) \in \boldsymbol{C}, \boldsymbol{C} \in \mathcal{A}(n)\}$ where $\mathcal{A}_n = (\mathcal{A}^1(n) \dots \mathcal{A}^D(n))$ We will bound the probability that DEX chooses a suboptimal action for clinic in an exploitation phase when context vector is in the set of active hypercubes $\boldsymbol{C}$ for any $\boldsymbol{C}$, and then use this to bound the expected number of times a suboptimal action is chosen by clinic for its patients in exploitation steps using DEX. Recall that reward loss in every step in which a suboptimal action is chosen can be at most 2.

Let $\mathcal{V}_{f,\boldsymbol{C}}(n)$ be the event that a suboptimal expert f is chosen for the set of hypercubes $\boldsymbol{C}$ by clinic for its nth patient. For $f \in \mathcal{F}$, let $\mathcal{E}_{F,\boldsymbol{C}}(n)$ be the set of rewards observed by clinic from expert f for its own patients whose context vectors is in the active set $\boldsymbol{C}$ up to the nth patient. For each set of hypercubes $\boldsymbol{C} = (C^1, \dots, C^D)$, let $f_n^*(\boldsymbol{C}) \in \mathcal{F}$ be the expert which is optimal for the center context of the type-d hypercube which has the highest expected diagnostic reward among all types of contexts for $\boldsymbol{C}$ for the patient n, and let $d_{\text{rel}}(\boldsymbol{C})$ be the type of the context for which diagnostic action $f_n^*(\boldsymbol{C})$ has the highest expected reward.

We generate two different artificial i.i.d. processes to bound the probabilities related to deviation of sample mean reward estimates $\bar{r}^d_{f,C^d}(n)$, $f \in \mathcal{F}$, $d \in \mathcal{D}$ from the expected rewards, which will be used to bound the probability of choosing a suboptimal action. The first one is the *best* process in which rewards are generated according to a bounded i.i.d. process with expected reward $\overline{\mu}^d_{f,C^d}$, the other one is the *worst* process in which the rewards are generated according to a bounded i.i.d. process with expected reward $\underline{\mu}^d_{f,C^d}$. Let $\bar{r}^{b,d}_{f,C^d}(n)$ denote the sample mean of the n samples from the best process and $\bar{r}^{w,d}_{f,C^d}(n)$ denote the sample mean of the n samples from the worst process. We have for any $f \in \mathcal{L}_{\boldsymbol{C}}(n)$

$$
\begin{aligned}
&P\left(\mathcal{V}_{f,\boldsymbol{C}}(n), \mathcal{W}_{\boldsymbol{C}}(n)\right) \\
&\leq P\left(\max_{d\in\mathcal{D}} \bar{r}^{b,d}_{f,C^d}(T^d_{f,C^d}(n)) \geq \overline{\mu}_{f,\boldsymbol{C}} + H_n, \mathcal{W}_{\boldsymbol{C}}(n)\right) \\
&+ P\left(\max_{d\in\mathcal{D}} \bar{r}^{b,d}_{f,C^d}(T^d_{f,C^d}(n)) \geq \bar{r}^{w,d_{\text{rel}}(\boldsymbol{C})}_{f_n^*(\boldsymbol{C}),C^{d_{\text{rel}}(\boldsymbol{C})}}(T^{d_{\text{rel}}(\boldsymbol{C})}_{f_n^*(\boldsymbol{C}),C^{d_{\text{rel}}(\boldsymbol{C})}}(n))\right. \\
&, \max_{d\in\mathcal{D}} \bar{r}^{b,d}_{k,C^d}(T^d_{k,C^d}(n)) < \overline{\mu}_{k,\boldsymbol{C}} + L2^{-l_{\max}(\boldsymbol{C})\alpha} \\
&+H_n, \bar{r}^{w,d_{\text{rel}}(\boldsymbol{C})}_{f_n^*(\boldsymbol{C}),C^{d_{\text{rel}}(\boldsymbol{C})}}(T^{d_{\text{rel}}(\boldsymbol{C})}_{f_n^*(\boldsymbol{C}),C^{d_{\text{rel}}(\boldsymbol{C})}}(n)) \\
&\left. > \underline{\mu}_{f_n^*(\boldsymbol{C}),\boldsymbol{C}} - L2^{-l_{\max}(\boldsymbol{C})\alpha} - H_n, \mathcal{W}_{\boldsymbol{C}}(n)\right) \qquad (1)
\end{aligned}
$$

$$
\begin{aligned}
&+ P\left(\max_{d\in\mathcal{D}} \bar{r}^{w,d^*(\boldsymbol{C})}_{f_n^*(\boldsymbol{C}),C^d}(T^d_{f_n^*(\boldsymbol{C}),C^d}(n)) \leq \underline{\mu}_{f_n^*(\boldsymbol{C}),\boldsymbol{C}} - H_n\right. \\
&\left., \mathcal{W}_{\boldsymbol{C}}(n)\right),
\end{aligned}
$$

where $H_n > 0$. In order to make the probability in (1) equal to 0, we need

$$2H_n \leq (B-2L)2^{-l_{\max}(\boldsymbol{C})\alpha}. \qquad (2)$$

By Lemma 1, (2) holds when

$$2H_n \leq (B-2L)2^{-\alpha}n^{-\alpha/p}. \qquad (3)$$

For $H_n = n^{z/2}$, $z \geq 2\alpha/p$ and $B = 2/(2^{-\alpha}) + 2L$, (3) holds by which (1) is equal to zero. Also by using a Chernoff-Hoeffding bound we can show that

$$P\left(\max_{d\in\mathcal{D}} \bar{r}^{b,d}_{f,C^d}(T^d_{k,C^d}(n)) \geq \overline{\mu}_{f,\boldsymbol{C}} + H_n, \mathcal{W}_{\boldsymbol{C}}(n)\right) \leq D/n^2,$$

and

$$
\begin{aligned}
&+ P\left(\max_{d\in\mathcal{D}} \bar{r}^{w,d}_{f_n^*(\boldsymbol{C}),C^d}(T^d_{f_n^*(\boldsymbol{C}),C^d}(n)) \leq \underline{\mu}_{f_n^*(\boldsymbol{C}),\boldsymbol{C}} - H_n\right. \\
&\left., \mathcal{W}_{\boldsymbol{C}}(n)\right) \leq D/n^2.
\end{aligned}
$$

$P\left(\mathcal{V}_{f,\boldsymbol{C}}(n), \mathcal{W}_{\boldsymbol{C}}(n)\right) \leq 2D/n^2$, for $f \in \mathcal{F}$ and We get the final bound by summing these probabilities from $n=1$ to N. ☐

Lemma 4. ***One-step regret due to near-optimal experts for a set of hypercubes.*** *For any set of hypercubes $\boldsymbol{C}$, the one-step regret of clinic from choosing one of its near optimal experts is bounded above by $B2^{-l_{\max}(\boldsymbol{C})\alpha}$.*

Proof. For the nth patient of clinic if $\boldsymbol{x}(n) \in \boldsymbol{C} \in \mathcal{A}(n)$,The per-patient regret of any near optimal expert $f \in \mathcal{F}$ is bounded by $BL2^{-l_{\max}(\boldsymbol{C})\alpha}$ by its definition. ☐

Lemma 5. *The regret due to near- optimal experts is $N^{\frac{1+p-\alpha}{1+p}}$*

Proof. For the nth patient of clinic, for the set of active hypercubes $\boldsymbol{C}(n)$ that the patient's context vector belongs to, $l_{\max}(\boldsymbol{C}(n))$ is at least the level of the active hypercube $x^d(n) \in C^d(n)$ for some type-d context. Since a near optimal action's regret for the nth patient is upper bounded by $2B2^{-l_{\max}(\boldsymbol{C}(n))\alpha}$, the total regret due to near optimal actions up to the Nth patient is upper bounded by $2B\sum_{n=1}^{N} 2^{-l_{\max}(\boldsymbol{C}(n))\alpha} \leq 2B\sum_{n=1}^{N} 2^{-l(C^d(n))\alpha}$.

Let $l_{\max}$ be the maximum level type-d hypercube when type-d contexts of the first N patients are uniformly distributed. We must have $A\sum_{l=1}^{l_{\max,u}-1} 2^l 2^{pl} < N$, otherwise the highest level hypercube for the Nth patient will be $l_{\max,u}-1$. Solving this equation for $l_{\max,u}$, we get $l_{\max,u} < 1+\log_2(N)/(1+p)$. $\sum_{n=1}^{N} 2^{-l(C^d(n))\alpha}$ takes its greatest value when type-d context up to the Nth patient is uniformly distributed in $\mathcal{X}_d$. Therefore we have $\sum_{n=1}^{N} 2^{-l(C^d(n))\alpha} \leq \frac{A2^{2(1+p-\alpha)}}{2^{1+p-\alpha}-1} N^{\frac{1+p-\alpha}{1+p}}$.

□

7.2. Proof of Theorem 1 and 2

Proof. For each hypercube of each type-d context, the regret due to trainings and explorations is bounded by Lemma 1. It can be shown that for each type-d context there can be at most $4N^{1/(1+p)}$ hypercubes that are activated up to the Nth patient. Using this we get a $O(N^{z+1/(1+p)}\log N)$ upper bound on the regret due to explorations and trainings for a type-d context. Then we sum over all types of contexts $d \in \mathcal{D}$. We show in Lemma 4 that the regret due to near optimal expert recommendations in exploitation phases is $O(N^{\frac{1+p-\alpha}{1+p}})$. In order to balance the order of regret (in the number of patients) due to explorations, trainings and near optimal action selections in exploitations, while at the same time minimizing the number of explorations and trainings, we set $z = 2\alpha/p$, and $p = \frac{3\alpha+\sqrt{9\alpha^2+8\alpha}}{2}$. Notice that we do not need to balance the order of regret due to suboptimal expert recommendations since its order is always less than the order of trainings and explorations. We get the final result by summing these two terms together with the regret due to suboptimal action selections in exploitation phases which is given in Lemma 2.

Theorem 2 directly follows from Lemma 3. □

8. REFERENCES

[1] Casey C Bennett and Kris Hauser, "Artificial intelligence framework for simulating clinical decision-making: A Markov decision process approach," *Artificial Intelligence in Medicine*, vol. 57, no. 1, pp. 9–19, 2013.

[2] Joann G Elmore, Diana L Miglioretti, Lisa M Reisch, Mary B Barton, William Kreuter, Cindy L Christiansen, and Suzanne W Fletcher, "Screening mammograms by community radiologists: variability in false-positive rates," *Journal of the National Cancer Institute*, vol. 94, no. 18, pp. 1373–1380, 2002.

[3] Metin N Gurcan, Berkman Sahiner, Nicholas Petrick, Heang-Ping Chan, Ella A Kazerooni, Philip N Cascade, and Lubomir Hadjiiski, "Lung nodule detection on thoracic computed tomography images: preliminary evaluation of a computer-aided diagnosis system," *Medical Physics*, vol. 29, no. 11, pp. 2552–2558, 2002.

[4] Jinshan Tang, Rangaraj M Rangayyan, Jun Xu, Issam El Naqa, and Yongyi Yang, "Computer-aided detection and diagnosis of breast cancer with mammography: recent advances," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 13, no. 2, pp. 236–251, 2009.

[5] Karthikeyan Ganesan, URajendra Acharya, Chua Kuang Chua, Lim Choo Min, KThomas Abraham, and K Ng, "Computer-aided breast cancer detection using mammograms: a review," *Biomedical Engineering, IEEE Reviews in*, vol. 6, pp. 77–98, 2013.

[6] Mostafa Fathi Ganji and Mohammad Saniee Abadeh, "A fuzzy classification system based on ant colony optimization for diabetes disease diagnosis," *Expert Systems with Applications*, vol. 38, no. 12, pp. 14650–14659, 2011.

[7] Amit X Garg, Neill KJ Adhikari, Heather McDonald, M Patricia Rosas-Arellano, PJ Devereaux, Joseph Beyene, Justina Sam, and R Brian Haynes, "Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review," *JAMA*, vol. 293, no. 10, pp. 1223–1238, 2005.

[8] Ashly D Black, Josip Car, Claudia Pagliari, Chantelle Anandan, Kathrin Cresswell, Tomislav Bokun, Brian McKinstry, Rob Procter, Azeem Majeed, and Aziz Sheikh, "The impact of eHealth on the quality and safety of health care: a systematic overview," *PLoS Medicine*, vol. 8, no. 1, pp. e1000387, 2011.

[9] Bo Xin, Yoshinobu Kawahara, Yizhou Wang, and Wen Gao, "Efficient generalized fused lasso and its application to the diagnosis disease," in *Twenty-Eighth AAAI Conference on Artificial Intelligence*. AAAI, 2014.

[10] Martin Sewell, "Ensemble learning," *RN*, vol. 11, no. 02, 2008.

[11] Ethem Alpaydin, *Introduction to machine learning*, The MIT Press, 2004.

[12] Leo Breiman, "Bagging predictors," *Machine learning*, vol. 24, no. 2, pp. 123–140, 1996.

[13] Peter Bühlmann and Bin Yu, "Boosting with the l 2 loss: regression and classification," *Journal of the American Statistical Association*, vol. 98, no. 462, pp. 324–339, 2003.

[14] Aleksandar Lazarevic and Zoran Obradovic, "The distributed boosting algorithm," in *Proc. of the seventh ACM SIGKDD international conference on knowledge discovery and data mining*. ACM, 2001, pp. 311–316.

[15] John Langford and Tong Zhang, "The epoch-greedy algorithm for contextual multi-armed bandits," *Advances in Neural Information Processing Systems*, vol. 20, pp. 1096–1103, 2007.

[16] Aleksandrs Slivkins, "Contextual bandits with similarity information," in *24th Annual Conference on Learning Theory (COLT)*, 2011.

[17] Tyler Lu, David Pal, and Martin Pal, "Contextual multi-armed bandits," in *International Conference on Artificial Intelligence and Statistics*, 2010, pp. 485–492.

[18] Wei Chu, Lihong Li, Lev Reyzin, and Robert E Schapire, "Contextual bandits with linear payoff functions," in *Proc. of the International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2011.

[19] William Hsu, Ricky K Taira, Suzie El-Saden, Hooshang Kangarloo, and Alex AT Bui, "Context-based electronic health record: toward patient specific healthcare," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 16, no. 2, pp. 228–234, 2012.

[20] Onur Atan, William Hsu, Cem Tekin, and Mihaela van der Schaar, "A data-driven approach for matching clinical expertise to individual cases," http://medianetlab.ee.ucla.edu/papers/DEX.pdf.