Discover the Expert: Context-Adaptive Expert Selection for Medical Diagnosis

Cem Tekin, Onur Atan, Mihaela van der Schaar
Electrical Engineering Department, University of California, Los Angeles
Email: cmtkn@ucla.edu, oatan@ucla.edu, mihaela@ee.ucla.edu

***Abstract*—Medical diagnosis and treatment currently follow standardized processes. For instance, patients are assigned to medical experts (clinicians) based on either their availability or fixed assignment rules which solely consider the qualifications of the clinician and the symptoms of the patient. Such assignments can be highly inefficient in practice, since the expertise and experience of healthcare providers varies significantly. Each clinician's level of expertise is unknown a priori and can only be estimated based on the accuracy of his/her past diagnostic decisions. Moreover, the level of expertise often depends on the "context" of the patient (e.g. health condition, family history). In this paper we propose an expert selection system that learns online the best expert to assign to each patient depending on the context of the patient. In general, the context can include an enormous number and variety of information related to the patient's health condition, age, gender, previous drug doses, etc., but the most relevant information is embedded in only a few contexts. If these most relevant contexts were known in advance, learning would be relatively simple - but they are not. Moreover, the relevant contexts may be different for different health conditions. To address these challenges, we develop a new class of algorithms aimed at discovering the most relevant contexts and the best clinic and expert to use to make a diagnosis given a patient's contexts. We prove that as the number of patients grows, the proposed context-adaptive algorithm will discover the optimal 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, which can be taken into account by the primary care physician before making the final decision. While our algorithm is general and can be applied in numerous medical scenarios, we illustrate its functionality and performance by applying it to a real-world breast cancer diagnosis dataset. Finally, while the application we consider in this paper is medical diagnosis, our proposed algorithm can be applied in other semantic computing systems, in which contexts are used to learn and select experts including personalized education, personalized recommendations, and business intelligence.**

***Index Terms*—Semantic computing, context-adaptive learning, clinical decision support systems, healthcare informatics, distributed multi-user learning, contextual bandits.**

I. Introduction

One of the most important applications of semantic computing [1] is healthcare informatics [2]. The development of healthcare informatics tools and decision support systems is vital, since recent studies show that standard clinical practice often fails to fit the patient [3]. The main reasons for this are the diverse types of patients and their health states, diverse level of expertise exhibited by healthcare professionals as well as the diversity of the tests and equipment available at different health clinics. For instance, for breast cancer diagnosis, it is observed that younger and inexperienced radiologists have higher false-positive rates than experienced radiologists when detecting breast cancer from radiological images, and the detection accuracy over different healthcare providers ranges from 2.6% to 15.9% [4]. In order to compensate for these differences and provide uniformly good healthcare, *Clinical Decision Support Systems* (CDSSs) have been integrated as part of numerous clinical decision making processes including diagnostic decisions for lung cancer [5], breast cancer [6], [7] and diabetes [8]. However, most existing CDSSs are dedicated to specific deployments, specific health professionals and specific areas of expertise. These systems are trained using past clinical data, and then tested on small patient populations. Although recent surveys show that CDSSs have improved the accuracy of clinical decision making process for some applications [9], their effectiveness cannot be validated for other applications [10]. For instance, in [10] a detailed empirical analysis of different CDSSs was carried out and the analysis showed that the cost and health risks of deploying CDSSs often outweighs the benefits. It is also shown that 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, the ease of use of the CDSS by healthcare professionals, etc. In conclusion, one of the key challenges in health clinics is how to decide when to rely on the decisions of a CDSS and which CDSS to use, when to follow the advice of a human expert and which expert to choose and possibly when to select another clinic to make the diagnosis decision.

Unlike existing prior work in health informatics which considers the problem of designing various CDSSs, in this paper we study the broader problem of integrating CDSSs and human experts over multiple clinics in an efficient way to improve the diagnostic decision making process. We assume that the diagnostic accuracy of an expert (either human or CDSS) depends on the *context* of the patient for which the decision is made. This context is all the information pertaining to the patient under consideration that can be utilized in the decision making process. For instance, in breast cancer diagnosis context includes patient profile, breast density, assessment history, characteristics of the opposite breast, modalities, etc., or in general, electronic medical records can be used as

The material is based upon work funded by the US Air Force Research Laboratory (AFRL) under the DDAS program. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not reflect the views of AFRL.

The authors would like to acknowledge prof. William Hsu (Radiology Department, UCLA) for numerous useful discussions and guidance and for providing us the data set used in the illustrative results section.

the context [11]. Since the context of a patient has many dimensions, learning the diagnostic accuracy of an expert suffers from the curse of dimensionality. The methodology we propose in this paper learns the *most relevant* context(s) pertinent to the current health condition of the patient and use it/them the estimate the level of expertise exhibited by the expert. The level of expertise is defined based on the accuracy of their diagnostic.

In addition to specific information about the patient, the context vector can contain *semantic* information related to type of diagnosis to be made, testing and imaging equipment, etc. This information is exploited when learning the context-specific diagnostic accuracy of the experts. For instance, while the diagnostic accuracy of human experts is high for most breast images, there exist a small number of breast images for which it is particularly difficult to discern the abnormality [12]. For these cases, it may be better to use a CDSS. Moreover, different clinics have healthcare professionals with different expertise and some of these clinics may have access to CDSSs from different manufacturers and of different types while some others just rely on human experts. In our proposed system, these clinics can cooperate with each other to improve diagnostic accuracy by learning the contextual specializations of the other clinics (see Fig. 1). For instance, a rural clinic may have only a primary care physician (PCP), a registered nurse and equipment, but no specialist or CDSSs and may be able to request information from a more established hospital which implements multiple CDSSs and has several experienced human experts. Moreover, privacy policies may not allow a clinic to send the entire patient information to another clinic [13] (only anonymized patient information can be transferred among clinics). In addition, it will not be possible to know the exact decision making mechanism used by the other clinic (e.g., the rule that a CDSS uses to output the diagnosis recommendation or the mechanism that a human expert uses when making a diagnosis recommendation). Hence, based on the context of the patient, each clinic learns (i) whether it should rely on its own experts or request another clinic to make diagnostic decisions, and (ii) if it relies on its own experts, which expert it should assign the task of deciding the diagnostic, such that the reward (gain) obtained by selecting that particular expert is maximized. The reward for a particular diagnostic decision can be defined as the diagnostic accuracy minus the incurred cost (e.g. delay, money, etc.).

Our proposed system learns online, meaning that its expert selection strategy is updated every time after the true health state of a patient is revealed. (Note that at times the true health state is not revealed immediately.) Based on this feedback, the expertise, i.e., the diagnostic accuracy, of the chosen expert is updated.

We model this problem as a distributed context-adaptive online learning problem. Each clinic decides on what diagnostic action to take based on the history of its own patient arrivals, patient arrivals to other clinics that requested a diagnostic action from the current clinic, and the success rate of each diagnostic action. In this way, each clinic is able to identify which experts make accurate decisions for patients with specific contexts. As a side result, clinics that use CDSSs learn whether the deployed CDSSs are beneficial for patients given their contexts. The main contributions of this paper are summarized below:

- We propose a novel algorithm that learns the diagnosis accuracy of different clinics, human experts and CDSSs. The algorithm is distributed and uses the context of each patient to learn the best expert, i.e. the expert who has the highest reward, which trades off the diagnosis accuracy and the costs of making the diagnose.
- Since the set of contexts is large, the proposed algorithm discovers the most relevant context or set of contexts that it should use for selecting an expert, i.e., experts are selected based on the relevant characteristics (contexts) of the patients. Hence, it learns fast, and the learning speed is independent of the dimension of the context vector. As a side benefit, the algorithm also learns what experts are good at diagnosing patients exhibiting specific contexts (e.g. specific age groups, genders, health states, etc.).
- The proposed algorithm promotes cooperation among clinics. A clinic learns the expertise of its own human experts and CDSSs not only from its own patients but also from the patients of other clinics who request diagnosis recommendation from that clinic. Moreover, each clinic benefits from the expertise of other clinics by asking them to make diagnosis recommendations for their patients.
- The proposed algorithm learns the optimal experts to use very fast: we formally determine performance bounds for the proposed algorithm which quantify how fast it can learn the optimal expert given the number of patients that have been treated so far.
- For each patient, the algorithm provides confidence bounds on the accuracy of the selected diagnostic action, and guarantees that the expert chosen for the diagnosis is the best expert or a near-optimal expert with a high probability.
- Patient safety is one of the most important aspects of healthcare. Our proposed algorithm merely makes diagnosis recommendations to the PCP or the clinician in charge for the care of the patient. The final decision of following the diagnosis recommendation or assigning another expert for another diagnosis recommendation is up to the PCP/clinician. Such additional layers of safety do not change the operation of our algorithm and allow for personalization by the PCP/clinician.

The remainder of the paper is organized as follows. In Section II, we describe the related work. In Section III, we formalize the problem and we present the proposed distributed, context-adaptive online algorithm which learns the best expert for diagnosing a patient based on his/her contextual information in Section IV. We illustrate the proposed system using a real-world breast cancer diagnostic data in Section V. Finally, the concluding remarks are given in Section VI.

II. Related Work

We categorize the related work into three key areas: work related to semantic computing, work related to CDSSs, and work related to data mining and online learning.

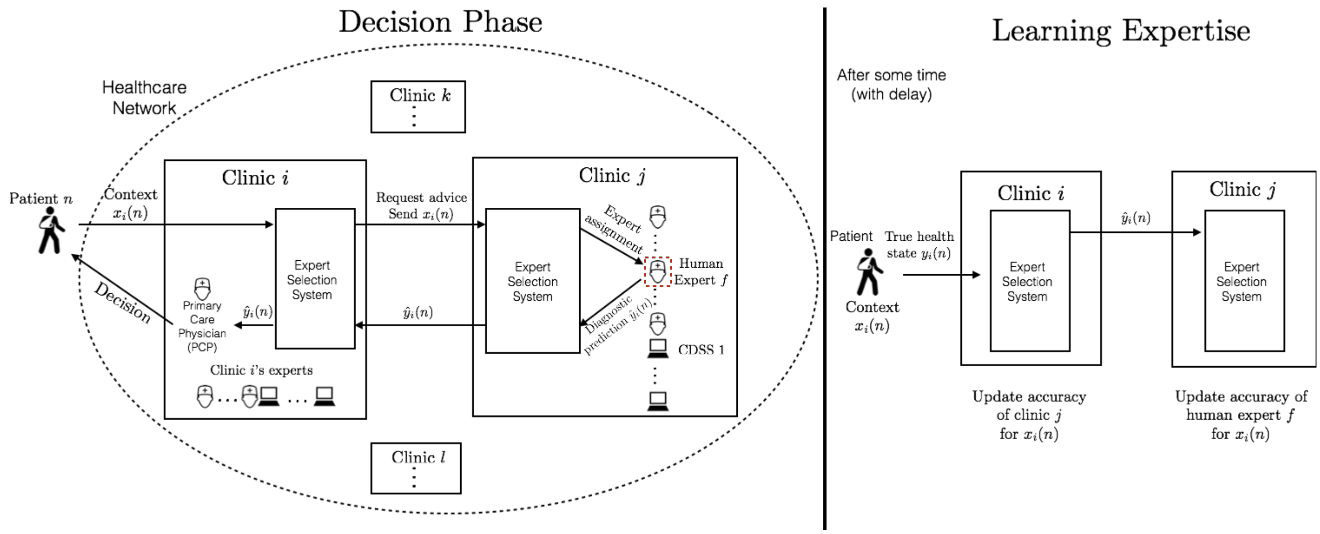

Fig. 1. Operation of the proposed system for clinic i. In this example, the diagnostic decision for the patient with context $x_i(n)$ is made by a human expert from clinic j. Then, after some time the true health state $y_i(n)$ of the patient is revealed. Based on this clinic i updates the diagnostic accuracy of clinic j for that context, while clinic j updates the diagnostic accuracy of its expert f.

A. Semantic computing

Semantic computing focuses on computing based on semantics ("context", "meaning", "intention") and it addresses all types of resources including data, document, tool, device, process and people [1]. Within the area of semantic computing, rule-based reasoning systems [14], [15] have emerged which deploy a database of the facts that are known about the problem currently being solved, and a decision engine which combines rules with the data to produce predictions. In these systems the decision rule is developed by a group of human experts, and rules are updated over time based on their effectiveness. Our proposed methodology fits within the class of semantic-based reasoning systems. However, in contrast to the existing work, we consider multiple experts, each adopting its own decision rule. Moreover, how well a specific decision rule (diagnostic) performs when applied to a patient, characterized by a specific context, is not known a priori. Hence, in this work we are interested in developing a rigorous and efficient methodology for learning how to select the expert adopting the best decision rule (diagnostic) can be learned for each patient.

B. Clinical decision support systems

Although we propose a method to integrate CDSSs and human experts among multiple clinics and learn their expertise to maximize the efficiency of the clinical diagnosis system, most of the prior work focused on designing a specific CDSSs for various diagnostic applications such as lung cancer [5], breast cancer [6], [7] and diabetes [8]. Our expert selection method supports and validates the efficiency of such CDSSs based on the characteristics of the patient population. Specifically, our method discovers the accuracy of a CDSS for groups of patients based on their contexts, and this information can be used by clinicians to improve the CDSS or discontinue its use. Moreover, our system allows choosing between a CDSS or a human expert. Hence based on the preferences of the PCP additional features such as "Choose the expert only among the CDSSs", "Choose the expert only among human experts" or "Choose the expert only from a specific clinic" can easily be integrated to the system.

C. Data mining and learning

Since contexts of the patients are a form of data, and since the patients are arriving sequentially over time, there is a close relationship between our proposed methodology and online stream mining. Most of the prior work in online stream mining provides algorithms which are asymptotically converging to an optimal or locally-optimal solution without providing any rates of convergence. On the contrary, we do not only prove convergence results, but we are also able to explicitly characterize the performance loss incurred at each time slot (for each patient) with respect to the optimal solution.

Some of the existing solutions (including [16]–[20]) propose ensemble learning techniques. In our work we only consider choosing the best expert (initially unknown) for each context. This is especially important in resource constrained scenarios like healthcare informatics, where the human resources are limited either in terms of the number of experts that are making diagnostic decisions or the number of healthcare personnel that acts as an interface between the patient and the CDSS. We provide a detailed comparison to our work in Table I. As seen from Table I, our proposed system is context-adaptive, distributed, outputs confidence bounds, and provides an explicit rate of convergence to the optimal expert selection strategy as the number of patients grows.

In addition to the problems in data mining, our methods can be applied to any problem that can be formulated as a distributed contextual bandit problem. Contextual bandits have been studied before in [21]–[24] and other works in a single agent setting. However our work is very different from these because (i) we consider decentralized agents (clinics) who can learn to cooperate with each other, (ii) the set of available (diagnostic) actions and the context arrivals to the agents can be very different for each agent, (iii) instead of learning to take the best action considering the entire D-dimensional context vector, an agent learns to take the marginally best action by independently considering each D types of contexts, hence learning is much faster. It is known that the performance of existing learning algorithms depend on the dimension of the context space [22]. Due to its context-adaptive property, the convergence speed of the algorithm we propose in this paper is independent of the dimension of the context space.

III. Problem Formulation

The system model is shown in Fig. 2 and 3. There are M clinics (learners) which are indexed by the set $\mathcal{M} := \{1, 2, \ldots, M\}$. The set of experts clinic i has is $\mathcal{F}_i$. As we discussed an expert can either be a human expert or a CDSS. The set of all experts is $\mathcal{F} = \cup_{i \in \mathcal{M}} \mathcal{F}_i$. Let $\mathcal{M}_{-i} := \mathcal{M} - \{i\}$ be the set of clinics clinic i can choose from to send its patient's context for diagnosis. The *diagnostic action set*[1] of clinic i is $\mathcal{K}_i := \mathcal{F}_i \cup \mathcal{M}_{-i}$. Throughout the paper we use index f to denote an element of $\mathcal{F}$, j to denote clinics in $\mathcal{M}_{-i}$, and k to denote an element of $\mathcal{K}_i$.

For each patient $n = 1, 2, \ldots, N$, the following events happen sequentially: (i) The nth patient with a D-dimensional context vector $\boldsymbol{x}_i(n) = (x_i^1(n), \ldots, x_i^D(n))$ arrives to clinic $i \in \mathcal{M}$, where $x_i^d(n) \in \mathcal{X}_d$ for $d \in \mathcal{D} := \{1, \ldots, D\}$ and $\mathcal{X}_d$ is the set of type-d contexts, and $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_D$ is the context space,[2] (ii) each clinic i assigns one of its own experts or another clinic to recommend a diagnosis $\hat{y}_i(n) \in \mathcal{Y}$ for the patient n, 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), (iii) after some delay, the true health state of patient $y_i(n) \in \mathcal{Y}$ is revealed only to the clinic i where the patient has arrived,[3] (iv) if another clinic provided the diagnosis for that patient, then the clinic where the patient arrived passes the true health state of the patient to that clinic.

A. Context, Diagnosis, Diagnostic Action Accuricies

For each patient n, the context vector $\boldsymbol{x}_i(n)$ and true health state of the patient $y_i(n)$ are assumed to be drawn from an (unknown) joint distribution J over $\mathcal{X} \times \mathcal{Y}$ independently from the other patients. We do not require this draw to be independent among the clinics/learners. Since the context vector of the patient $\boldsymbol{x}_i(n)$ is revealed to the clinic i, depending on J, there exists a conditional distribution $G_{\boldsymbol{x}_i(n)}$ over $\mathcal{Y}$. Similarly, depending on J, there is a marginal distribution H over $\mathcal{X}$ from which contexts are drawn. Given context vector $\boldsymbol{x}$, let $\pi_f(\boldsymbol{x}) = \int_{y \in \mathcal{Y}} \mathrm{I}(f(\boldsymbol{x}_i(n)) = y_i(n)) dG_{\boldsymbol{x}}(y)$ be the *joint accuracy* (or simply, accuracy) of expert $f \in \mathcal{F}$, where $f(\boldsymbol{x}_i(n))$ is the diagnosis recommendation of expert f for context vector $\boldsymbol{x}_i(n)$.[4] The diagnostic rule used by expert f, i.e., $f(\cdot)$ is allowed to be deterministic or random. $\mathrm{I}(\cdot)$ is the indicator function which is equal to 1 if the statement inside is true and 0 otherwise, and the expectation $\mathrm{E}[\cdot]$ is taken with respect to distribution $G_{\boldsymbol{x}}$. Let $\boldsymbol{x}^{-d} := (x^1, \ldots, x^{d-1}, x^{d+1}, \ldots, x^D)$ and $((\boldsymbol{x}')^{-d}, x^d) = (x'^1, \ldots, x'^{d-1}, x^d, x'^{d+1}, \ldots, x'^D)$. Then, the *marginal accuracy* of expert f based on type-d context is defined as

$$\pi_f^d(x^d) := \int_{(\boldsymbol{x}')^{-d}} \pi_f((\boldsymbol{x}')^{-d}, x^d) dH((\boldsymbol{x}'^{-d}), x^d).$$

We say that the problem has *the similarity property* when each expert has similar marginal accuracies for similar contexts.

Definition 1: **Similarity Property.** If the joint distribution J over $\mathcal{Y} \times \mathcal{X}$ is such that for each $f \in \mathcal{F}$ and $d \in \mathcal{D}$, there exists a minimum $\alpha > 0$ and a minimum $L > 0$, such that for all $x^d, (x')^d \in \mathcal{X}_d$, we have $|\pi_f^d(x^d) - \pi_f^d((x')^d)| \leq L|x^d - (x')^d|^\alpha$, then we call J *a distribution with similarity*.

Although, our model assumes a continuous context space, our algorithms will also work when the context space is discrete. Note that Definition 1 does not require the context space to be continuous. We assume that α is known by the clinics, while L does not need to be known. However, our algorithms can be combined with estimation methods for α. In reality, the knowledge of α is not required for our algorithms to run, however an estimate of α should be given as an input. If the estimate $\hat{\alpha}$ is chosen conservatively such that $\hat{\alpha} < \alpha$, the performance bounds we prove for our algorithm (Theorem 1 and Corollary 1) will hold with α replaced by $\hat{\alpha}$.

B. Unknowns, experts and diagnostic rewards

In our problem, the unknowns for clinic i are (i) $\mathcal{F}_j$, $j \in \mathcal{M}_{-i}$, (ii) J, H, $G_{\boldsymbol{x}}$, $\boldsymbol{x} \in \mathcal{X}$, (iii) $\pi_f(\boldsymbol{x})$, $f \in \mathcal{F}_i$, $\boldsymbol{x} \in \mathcal{X}$, (iv) $\pi_f^d(x^d)$, $f \in \mathcal{F}_i$, $x^d \in \mathcal{X}_d$, $d \in 1, \ldots, D$.

On the other hand, clinic i knows (i) the experts in $\mathcal{F}_i$ and costs of calling them,[5] (ii) the set of other clinics $\mathcal{M}_{-i}$ and costs of calling them, (iii) and an upper bound on the number of experts that each clinic has, i.e., $F_{\max} \geq |\mathcal{F}_j|$,[6] for all $j \in \mathcal{M}_{-i}$.

For each patient n, clinic i can either assign one of its experts or forward the patient's context to another clinic to have him/her diagnosed. We assume that for clinic i, assigning each expert $f \in \mathcal{F}_i$ incurs a cost $c_f^i \geq 0$. We assume that whenever the patient's context is sent to another clinic $j \in \mathcal{M}_{-i}$ a cost of c_j^i is incurred by clinic i.[7] This cost can be the delay and/or monetary costs associated with the forwarding action. When the diagnostic action $k \in \mathcal{K}_i$ is chosen for the

[1] In sequential online learning literature [29], [30], an action is also called an arm (or an alternative).

[2] 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. In our analysis, we will assume that $\mathcal{X}_d = [0, 1]$ for all $d \in \mathcal{D}$. However, our algorithms will work and our results will hold even when the context space is discrete given that it is bounded. For instance, $\mathcal{X}_d$ can represent the set of normalized ages (formed for instance, by dividing the exact age of the patient with the maximum age 200). Then, every patient's normalized age will lie in $\mathcal{X}_d$.

[3] 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.

[4] Although for simplicity of exposition we assumed that the decision only depends on the context vector, the radiological image can also be a part of the information sent to the expert which is denoted by $s_i(n)$. Then assuming that this data is i.i.d. given a context vector, the decision rule can be extended as $f(\boldsymbol{x}_i(n), s_i(b))$, and the expert accuracy can be defined analogously.

[5] Cost can be monetary cost, delay, etc.

[6] For a set A, let $|A|$ denote the cardinality of that set.

[7] The cost for clinic i does not depend on the cost of the expert chosen by clinic j. Since the clinics are cooperative, j will obey the rules of the proposed algorithm when assigning an expert to diagnose clinic i's patient. We assume that when called by clinic i, clinic j will select an expert from $\mathcal{F}_j$, but not forward i's patient's context to another clinic.

	[18]–[20], [25]	[26], [27]	[28]	[22], [23]	This work
Message exchange	none	context	training residual	none	context (adaptively)
Learning approach	offline/online	offline	offline	Non-Bayesian online	Non-Bayesian online
Learning from other's contexts	N/A	no	no	no	yes
Using other's experts	no	all	all	no	sometimes-adaptively
Rate of convergence	no	no	no	yes - dimension dependent	yes - dimension independent
Context adaptive	no	no	no	no	yes
Confidence bounds	no	no	no	yes	yes

TABLE I
COMPARISON WITH RELATED WORK IN DATA MINING AND LEARNING.

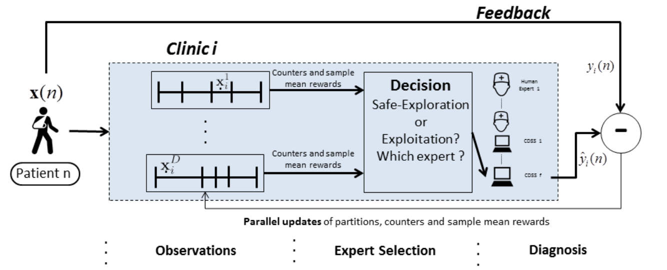

Fig. 2. Operation of clinic i for its nth patient when it chooses one of its own experts.

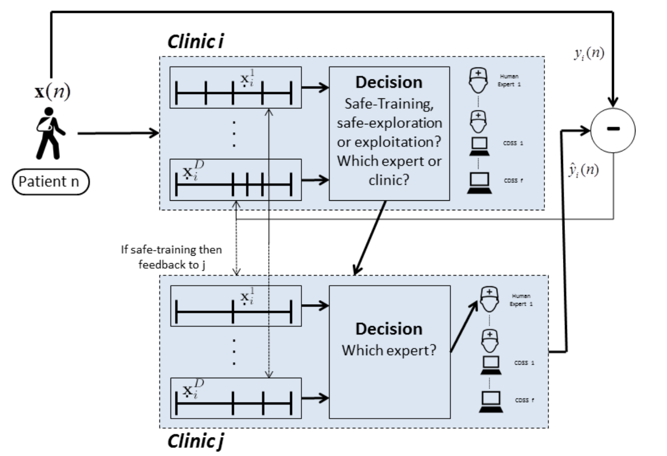

Fig. 3. Operation of clinic i for its nth patient when it chooses clinic j.

nth patient of clinic i, and the diagnosis recommendation $\hat{y}_i(n)$ is made, the reward is equal to $r_i(n) := \mathrm{I}(\hat{y}_i(n) = y_i(n)) - c^i_k$. This reward is observed only after the true health state $y_i(n)$ is revealed. Since the costs are bounded, without loss of generality we assume that costs are normalized, i.e., $c^i_k \in [0,1]$ for all $k \in \mathcal{K}_i$. The clinics are cooperative which implies that clinic $j \in \mathcal{M}_{-i}$ will return a diagnostic recommendation to i when called by i using its expert with the highest estimated diagnostic accuracy for i's context vector. Similarly, when called by $j \in \mathcal{M}_{-i}$, clinic i will return a diagnostic recommendation to j. In our theoretical analysis we do not consider the effect of this on i's learning rate; however, since our results hold for the case when other clinics are not forwarding their patient's context to i, they will also hold when other clinics forward the patient's context to i. Indeed, learning is faster for clinic i when other clinics ask clinic i for a diagnostic recommendation for their patients.

We assume that each expert produces a binary diagnostic recommendation,[8] thus $\mathcal{Y} = \{0,1\}$. For a clinic $j \in \mathcal{M}_{-i}$ its accuracy for a type-d context x^d is equal to the accuracy of its best expert, i.e.,

$$\pi^d_j(x^d) := \max_{f \in \mathcal{F}_j} \pi^d_f(x^d).$$

The goal of each learner i is to maximize its total expected reward. This corresponds to minimizing the regret with respect to the benchmark solution which we will define in the next subsection.

C. Diagnosis with Complete Information

Our benchmark when evaluating the performance of the learning algorithms is the solution which selects the diagnostic action in $\mathcal{K}_i$ with the highest marginal accuracy minus cost (i.e., reward) given the context vector $\boldsymbol{x}_i(n)$ for patient n. Specifically, the solution we compare against is given by

$$k^*_i(\boldsymbol{x}) := \arg\max_{k \in \mathcal{K}_i} \left(\max_{x^d \in \boldsymbol{x}} \pi^d_k(x^d) - c^i_k \right), \quad \forall \boldsymbol{x} \in \mathcal{X}.$$

Since calculating $k^*_i(\boldsymbol{x})$ requires knowledge of marginal expert accuracies only, we call $k^*_i(\boldsymbol{x})$ *the marginally best diagnostic action* given patient's context $\boldsymbol{x}$. Knowing this means that clinic i knows the expert in $\mathcal{F}$ that yields the highest diagnostic reward for each $x^d \in \mathcal{X}_d$, $d \in \mathcal{D}$. We call a policy that always acts according to this action an *optimal policy*.

D. The Regret of Learning

Simply, the regret is the loss incurred due to the unknown expertise. Regret of a learning algorithm which assigns an expert $a_i(n) \in \mathcal{K}_i$ for patient n in the clinic i based on its context vector $\boldsymbol{x}_i(n)$ and the past observations is defined as

$$R_i(N) := \sum_{n=1}^{N} \left(\pi_{k^*_i(\boldsymbol{x}_i(n))}(\boldsymbol{x}_i(n)) - c^i_{k^*_i(\boldsymbol{x}_i(n))} \right) - \mathrm{E}\left[\sum_{n=1}^{N} (\mathrm{I}(\hat{y}_i(n) = y_i(n)) - c^i_{a_i(n)}) \right],$$

where $\hat{y}_i(n)$ denotes the diagnostic recommendation of the expert or other clinic $a_i(n)$ assigned by clinic i to the patient n, $y_i(n)$ denotes the true health state of the patient n that arrived to clinic i. Regret gives the convergence rate of the total expected reward of the learning algorithm to the value of the benchmark solution $k^*_i(\boldsymbol{x})$, $\boldsymbol{x} \in \mathcal{X}$. Any algorithm whose regret is sublinear, i.e., $R_i(N) = O(N^\gamma)$ such that $\gamma < 1$, will converge to the benchmark solution in terms of the average reward.

IV. ADAPTIVELY LEARNING THE RELEVANT CONTEXTS

In this section we propose an online learning algorithm that achieves regret that is sublinear in the number of patients. We name our algorithm Learn the EXpert (LEX).

[8] In general we can assume that diagnostic recommendations belong to $\mathbb{R}$ and define diagnostic error as some other metric. Our results can be adapted to this case as well.

A. The LEX algorithm

The basic idea behind LEX is to learn the accuracies of different clinics and different experts by requesting diagnosis recommendations from them in a cost efficient way. There are three operation phases of LEX: exploitation, safe training and safe exploration. For any clinic and any patient LEX is only in one of these phases. In an exploitation phase, LEX is very confident about its expert selection decision. As we will show in Corollary 1, it is able provide confidence bounds on the probability that it selects the best expert among all possible experts and on the accuracy of the prediction made by the chosen expert. In the safe training phase, clinic i is not confident about how well some other clinic j knows its best expert for clinic i's patient. Hence, clinic i requests a diagnosis recommendation from clinic j which helps clinic j learn the accuracy of its own experts. In the safe exploration phase, clinic i is not confident about the accuracy of its diagnostic actions. It will choose a diagnostic action and receive a diagnosis recommendation to update the accuracy of the chosen diagnostic action (which is done after the true health state is revealed). Trainings and explorations are safe, which means that LEX alerts the clinician that is in charge of the patient that the diagnosis recommendation comes from an expert which may not be very reliable. Knowing this, the clinician may assign another expert or may choose to follow the recommendation based on his/her own expertise. This way the system learns, while the patient safety is not compromised. Whenever we refer to training and exploration, we mean safe training and safe exploration.

LEX 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 trained and 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 rewards of diagnostic actions in $\mathcal{K}_i$ for contexts that lie in those sets. For each patient, LEX chooses a diagnostic action adaptively based on the estimated marginal accuracy of the diagnostic action given the context vector. For the nth patient of clinic i, we call the type-d context for which the selected diagnostic action's marginal accuracy is maximized as the *main context* of that patient.

For each type-d context, LEX starts with a single hypercube which is the entire context space $\mathcal{X}_d$, then divides the space into finer regions and explores them as more patients with those contexts arrive. In this way, LEX focuses on parts of the context space in which there are large number of patient arrivals, and does this independently for each type of context of the patients.

The learning algorithm for clinic i should zoom into the regions of space with large number of context arrivals, but it should also persuade other clinics to *zoom* to the regions of the space where clinic i has a large number of context arrivals. Here, zooming means using past observations from a smaller region of context space to estimate the rewards of actions for a context. The pseudocode of LEX is given in Fig. 4, and the initialization, safe training, safe exploration and exploitation modules are given in Fig. 5 and Fig. 6.

Learn the Expert Algorithm (for clinic i):
Input: $D_1(n)$, $D_2(n)$, $D_3(n)$, p, A
Initialization: $\mathcal{A}_i^d = \{[0,1]\}$, $d \in \mathcal{D}$. $\mathcal{A}_i = \mathcal{A}_i^1 \times \ldots \times \mathcal{A}_i^D$. Run **Initialize**($\mathcal{A}_i$)
Notation: $\bar{\boldsymbol{r}}_k^i = (\bar{r}^{i,d}_{k,C^d(n)})_{d\in\mathcal{D}}$,
$\bar{\boldsymbol{r}}^i = (\bar{\boldsymbol{r}}_k^i)_{k\in\mathcal{K}_i}$,
l_C: level of hypercube C,
$\boldsymbol{T}_k^i = (T^{i,d}_{k,C^d(n)})_{d\in\mathcal{D}}$, $k \in \mathcal{K}_i$,
$\boldsymbol{T}^i = (\boldsymbol{T}_k^i)_{k\in\mathcal{K}_i}$.
while $n \geq 1$ **do**
  if $\exists d \in \mathcal{D}$ and $\exists k \in \mathcal{F}_i$ such that $T^{i,d}_{k,C^d(n)} \leq D_1(n)$ **then**
    Run **SafeExplore**(n, k, d, $T^{i,d}_{k,C^d(n)}$, $\bar{r}^{i,d}_{k,C^d(n)}$)
  else if $\exists d \in \mathcal{D}$ and $\exists k \in \mathcal{M}_{-i}$ such that $T^{i,d}_{1,k,C^d(n)} \leq D_2(n)$ **then**
    Run **SafeTrain**(n, k, d, $T^{i,d}_{1,k,C^d(n)}$)
  else if $\exists d \in \mathcal{D}$ and $\exists k \in \mathcal{M}_{-i}$ such that $T^{i,d}_{k,C^d(n)} \leq D_3(n)$ **then**
    Run **SafeExplore**(n, k, d, $T^{i,d}_{k,C^d(n)}$, $\bar{r}^{i,d}_{k,C^d(n)}$)
  else
    Run **Exploit**(n, $\boldsymbol{T}^i$, $\bar{\boldsymbol{r}}^i$, $\mathcal{K}_i$)
  end if
  $T^{i,d}_{C^d(n)} = T^{i,d}_{C^d(n)} + 1$
  for $d \in \mathcal{D}$ **do**
    if $T^{i,d}_{C^d(n)} \geq A2^{pl_{C^d(n)}}$ **then**
      Create 2 level $l_{C_d(n)} + 1$ child hypercubes denoted by $\mathcal{A}_{C^d(n)}$
      Run **Initialize**($\mathcal{A}_{C^d(n)}$)
      $\mathcal{A}_i = \mathcal{A}_i \cup \mathcal{A}_{C^d(n)} - C^d(n)$
    end if
  end for
  $n = n + 1$
end while

Fig. 4. Pseudocode of the LEX algorithm.

Initialize($\mathcal{A}$):
for $C \in \mathcal{A}$ **do**
  Set $T_C^{i,d} = 0$, $T^{i,d}_{k,C} = 0$, $\bar{r}^{i,d}_{k,C} = 0$ for $k \in \mathcal{K}_i$, $T^{i,d}_{1,k,C} = 0$ for $k \in \mathcal{M}_{-i}$.
end for

Fig. 5. Pseudocode of the initialization module.

For each type-d context, we call an interval $(a2^{-l}, (a+1)2^{-l}] \subset [0,1]$ a level l hypercube for $a = 1, \ldots, 2^l - 1$,[9] where l is an integer. Let $\mathcal{P}_l^d$ be the partition of type-d context space $[0,1]$ generated by level l hypercubes. Clearly, $|\mathcal{P}_l^d| = 2^l$. Let $\mathcal{P}^d := \cup_{l=0}^{\infty} \mathcal{P}_l^d$ denote the set of all possible hypercubes. Note that $\mathcal{P}_0^d$ contains only a single hypercube which is $\mathcal{X}_d$ itself. For each patient, LEX keeps for the clinic i a set of mutually exclusive hypercubes that cover the context space of each type-$d \in \mathcal{D}$ context. We call these hypercubes *active* hypercubes, and denote the set of active hypercubes for type-d context for patient n by $\mathcal{A}_i^d(n)$. Let $\mathcal{A}_i(n) := (\mathcal{A}_i^1(n), \ldots, \mathcal{A}_i^D(n))$. Clearly, we have $\cup_{C\in\mathcal{A}_i^d(n)} C = \mathcal{X}_d$. Denote the active hypercube that contains $x_i^d(n)$ by $C_i^d(n)$. Let $\boldsymbol{C}_i(n) := (C_i^1(n), \ldots, C_i^D(n))$ be the set of active hypercubes that contains $\boldsymbol{x}_i(n)$. The diagnostic action chosen by clinic i for patient n only depends on the diagnostic actions taken on previous context observations which are in $C_i^d(n)$ for some $d \in \mathcal{D}$. The number of such actions and observations

[9]The first level l hypercube is defined as $[0, 2^{-l}]$.

SafeTrain(n, k, d, $N^{i,d}_{1,k}$):

Select diagnostic action k.
Send type-d context $x_i^d(n)$ to clinic k.
Receive diagnosis recommendation $\hat{y}_i(n) = f_{k,n}(x_i^d(n)$ from clinic k, where $f_{k,n}$ is the expert that clinic k assigned to have the patient n diagnosed which is selected using $x_i^d(n)$.
A human expert (such as a PCP) evaluates $\hat{y}_i(n)$, and gives a final decision or asks other experts for opinion.
Receive true health state $y_i(n)$ (send this also to clinic k).
Compute reward $r_k(n) = \mathrm{I}(\hat{y}_i(n) = y_i(n)) - c_k^i$.
$T^{i,d}_{1,k,C^d(n)}{+}{+}$.

SafeExplore(n, k, d, $N^{i,d}_{k,C^d(n)}$, $\bar{r}^{i,d}_{k,C^d(n)}$):

Select diagnostic action k.
Receive diagnostic recommendation $\hat{y}_i(n)$.
A human expert (such as a PCP) evaluates $\hat{y}_i(n)$, and gives a final decision or asks other experts for opinion.
Receive true health state $y_i(n)$ (if $k \in \mathcal{M}_{-i}$, send this also to clinic k).
Compute diagnostic reward $r_k(n) = \mathrm{I}(\hat{y}_i(n) = y_i(n)) - c_k^i$.
$\bar{r}^{i,d}_{k,C^d(n)} = \frac{T^{i,d}_{k,C^d(n)} \bar{r}^{i,d}_{k,C^d(n)} + r_k(n)}{T^{i,d}_{k,C^d(n)} + 1}$.
$T^{i,d}_{k,C^d(n)}{+}{+}$.

Exploit(n, $\boldsymbol{N}^i$, $\hat{\boldsymbol{r}}^i$, $\mathcal{K}_i$):

Select diagnostic action $a_i(n) \in \arg\max_{j \in \mathcal{K}_i} \left(\max_{d \in \mathcal{D}} \bar{r}^{i,d}_{j,C^d(n)} \right)$.
Receive diagnostic recommendation $\hat{y}_i(n)$.
Receive true health state $y_i(n)$ (if $a_i(n) \in \mathcal{M}_{-i}$, send this also to clinic k).
Compute reward $r_{a_i(n)}(n) = \mathrm{I}(\hat{y}_i(n) = y_i(n)) - c^i_{a_i(n)}$.

Fig. 6. Pseudocode of the safe training, safe exploration and exploitation modules.

can be much larger than the number of previous actions and observations in $\boldsymbol{C}_i(n)$. This is because in order for an observation to be in $\boldsymbol{C}_i(n)$, it should be in all $C_i^d(n)$, $d \in \mathcal{D}$. Let $T_C^{i,d}(n)$ be the number of times type-d contexts have arrived to hypercube C of clinic i from the activation of C until patient n. Once activated, a level l hypercube C will stay active until the arrival of patient n such that $T_C^{i,d}(n) \geq A2^{pl}$, where $p > 0$ and $A > 0$ are parameters of LEX. After that, LEX will divide C into 2 level $l+1$ hypercubes.

For each expert in $\mathcal{F}_i$, LEX have a single (deterministic) control function $D_1(n)$ which controls when to do safe exploration or exploitation, while for each clinic in $\mathcal{M}_{-i}$, LEX have two (deterministic) control functions $D_2(n)$ and $D_3(n)$, where $D_2(n)$ controls when to do safe training, $D_3(n)$ controls when to do safe exploration or exploitation when there are enough trainings.

Marginal sample mean diagnostic action rewards, i.e., $\bar{r}^{i,d}_{k,C}(n)$, $k \in \mathcal{K}_i$, $C \in \mathcal{A}_i^d(n)$, are calculated only based on the rewards $r_i(n')$ collected through previous patients n' when k is selected due to a type-d context in C being under-explored (i.e., for patients n' such that $T^{i,d}_{k,C^d(n')} \leq D_1(n')$ for $k \in \mathcal{F}_i$ or $T^{i,d}_{k,C^d(n')} \leq D_3(n')$ for $k \in \mathcal{M}_{-i}$, and $C^d(n') = C$, $n' < n$ in the adaptive partition). This way, as the number of true health state observations increases, it is guaranteed that $\bar{r}^{i,d}_{k,C}(n)$ converges to a number very close to the true marginal expected reward action k for contexts in C. This, together with the adaptive partitioning guarantees that the regret remains sublinear in the number of patients.

In addition to the safe exploration phase, which allows clinics to build accurate estimates of the marginal expected diagnostic rewards of their experts for their own contexts, the safe training phase required for the diagnostic actions $k \in \mathcal{M}_{-i}$ serves the purpose of helping the clinics to build accurate estimates of the marginal expected diagnostic rewards for each other's patients. If clinic i forwards the patient's context to another clinic k, it cannot assign the expert that is selected by that clinic to have the patient diagnosed. If the estimated marginal accuracies of experts of clinic k are inaccurate, clinic i's estimate of clinic k's accuracy will be very different from the accuracy of clinic k's marginally optimal expert for clinic i's patient's context vector. Therefore, clinic i uses the rewards from clinic $k \in \mathcal{M}_{-i}$ to estimate the expected diagnostic reward of clinic k only if it believes that clinic k estimated the accuracies of its own experts accurately.

In order for clinic k to estimate the accuracies of its own experts accurately, if the number of patient arrivals to clinic k with contexts in set $C_i^d(n)$ is small, clinic i *trains* clinic k by sending its patient's type-d context to k, receiving back the diagnostic recommendation of the expert chosen by k and sending the true health state for that patient to k so that k can update the estimated accuracy of the expert (in $\mathcal{F}_k$) it had chosen for i. In order to do this, clinic i keeps two counters $T^{i,d}_{1,k,C}(n)$ and $T^{i,d}_{k,C}(n)$ for each $C \in \mathcal{A}^i_d(t)$, which are initially set to 0. If $T^{i,d}_{1,k,C}(n) \leq D_2(n)$, then clinic i trains clinic k by sending its patient's context $x_i^d(n)$, receiving a diagnostic recommendation from clinic k, and then sending the true health state $y_i(n)$ to clinic k so that clinic k can update the estimated marginal accuracy of the expert in $\mathcal{F}_k$ it had chosen to make a diagnostic recommendation for clinic i for type-d contexts. If $T^{i,d}_{1,k,C}(n) > D_2(n)$, for all $d \in \mathcal{D}$, this means that clinic k is trained enough for all types of contexts so it will almost always select the expert with the highest *marginal accuracy*, i.e., $\max_{d \in \mathcal{D}} \pi_f^d(x_i^d(n))$, $f \in \mathcal{F}_k$ when called by i.

To have sufficient observations from k before exploitation, i explores k when $T^{i,d}_{1,k,C}(n) > D_2(n)$ and $T^{i,d}_{k,C}(n) \leq D_3(n)$, and updates $T^{i,d}_{k,C}(n)$ and the sample mean marginal accuracy of clinic k, which is the ratio of the total number of correct diagnostic recommendations to the total number of diagnostic recommendations k has made for i for contexts in hypercube C. Let

$$\mathcal{S}^{i,d}_{C_i^d(n)}(n) := \Big\{ f \in \mathcal{F}_i \textbf{ : } T^{i,d}_{f,C_i^d(n)}(n) \leq D_1(n) \textbf{ or } j \in \mathcal{M}_{-i} \textbf{ : } T^{i,d}_{1,j,C_i^d(n)}(n) \leq D_2(n) \textbf{ or } T^{i,d}_{j,C_i^d(n)}(n) \leq D_3(n) \Big\},$$

and $\mathcal{S}^i_{\boldsymbol{C}_i(n)}(n) := \bigcup_{d \in \mathcal{D}} \mathcal{S}^{i,d}_{C_i^d(n)}(n)$. If $S^i_{\boldsymbol{C}_i(n)}(n) \neq \emptyset$ then LEX randomly selects an diagnostic action in $S^i_{\boldsymbol{C}_i(n)}(n)$ to train or explore, while if $S^i_{\boldsymbol{C}_i(n)}(n) = \emptyset$, LEX selects an diagnostic action in $\arg\max_{k \in \mathcal{K}_i} \left(\max_{d \in \mathcal{D}} \bar{r}^{i,d}_{k,C_i^d(n)}(n) \right)$ to exploit.

B. Analysis of the regret of LEX

In this subsection we analyze the regret of LEX and derive a sublinear upper bound on the regret, whose growth rate with the number of patients does not depend on D. We divide the regret $R_i(N)$ into three different terms. $R_i^e(N)$ is the regret due to trainings and exploitations by patients N, $R_i^s(N)$ is the regret due to selecting suboptimal diagnostic actions at exploitation steps by patients N, and $R_i^{ne}(N)$ is the regret due to selecting near-optimal actions in exploitation steps by patients N. Using the fact that trainings, explorations and exploitations are separated over time, and linearity of expectation operator, we get $R_i(n) = R_i^e(n) + R_i^s(n) + R_i^{ne}(n)$. In the following analysis, we will bound each part of the regret separately.

Lemma 1: **Regret of safe trainings and safe explorations in a hypercube.** Let $D_1(n) = D_3(n) = n^z \log n$ and $D_2(n) = F_{\max} n^z \log n$. For clinic i consider any level l hypercube for type-d contexts. The total regret in such a hypercube due to safe trainings and safe explorations up to the nth patient of clinic i is bounded by $O(MF_{\max} n^z \log n)$.

Proof: See Appendix A-B. ■

For clinic i let $\mu_k^d(x) := \pi_k^d(x) - c_k^i$, be the expected reward of diagnostic action $k \in \mathcal{K}_i$ for a patient with type-d context $x^d \in \mathcal{X}_d$. For each set of hypercubes $\boldsymbol{C} = (C^1, \ldots, C^D)$, let $k^*(\boldsymbol{C}) \in \mathcal{K}_i$ be the diagnostic action 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 $\boldsymbol{C}$, and let $d^*(\boldsymbol{C})$ be the type of the context for which diagnostic action $k^*(\boldsymbol{C})$ has the highest expected reward. Let $\overline{\mu}_{k,C^d}^d := \sup_{x \in C^d} \mu_k^d(x)$, $\underline{\mu}_{k,C^d}^d := \inf_{x \in C^d} \mu_k^d(x)$ $\overline{\mu}_{k,\boldsymbol{C}} := \max_{d \in \mathcal{D}} \overline{\mu}_{k,C^d}^d$, and $\underline{\mu}_{k,\boldsymbol{C}} := \max_{d \in \mathcal{D}} \underline{\mu}_{k,C^d}^d$, for $k \in \mathcal{K}_i$. When the set of active hypercubes of clinic i is $\boldsymbol{C}$, the set of suboptimal diagnostic actions is given by

$$\mathcal{L}_{\boldsymbol{C},B}^i := \left\{ k \in \mathcal{K}_i : \underline{\mu}_{k^*(\boldsymbol{C}),\boldsymbol{C}} - \overline{\mu}_{k,\boldsymbol{C}} > BL2^{-l_{\max}(\boldsymbol{C})\alpha} \right\},$$

where $B > 0$ is a constant and $l_{\max}(\boldsymbol{C})$ is the level of the highest level hypercube in $\boldsymbol{C}$. When the context vector of the patient is in $\boldsymbol{C}$, any diagnostic action that is not in $\mathcal{L}_{\boldsymbol{C},B}^i$ is a near-optimal diagnostic action.

Lemma 2: **Regret due to suboptimal diagnostic action selections.** Let $\mathcal{L}_{\boldsymbol{C},B}^i$, $B = 12/(L2^{-\alpha}) + 2$ denote the set of suboptimal diagnostic actions for set of hypercubes $\boldsymbol{C}$. When LEX is run with parameters $p > 0$, $2\alpha/p \leq z < 1$, $D_1(n) = D_3(n) = n^z \log n$ and $D_2(t) = F_{\max} n^z \log n$, the regret of clinic i due to choosing suboptimal diagnostic actions in $\mathcal{L}_{\boldsymbol{C}_i(n),B}^i$ for patients $1 \leq n \leq N$ in exploitation steps, i.e., $R_i^s(N)$, is bounded by $O(MF_{\max} N^{z/2})$.

Proof: See Appendix A-C. ■

Lemma 3: **Regret due to near-optimal clinics choosing suboptimal experts.** Let $\mathcal{L}_{\boldsymbol{C},B}^i$, $B = 12/(L2^{-\alpha}) + 2$ denote the set of suboptimal actions for set of hypercubes $\boldsymbol{C}$. When LEX is run with parameters $p > 0$, $2\alpha/p \leq z < 1$, $D_1(n) = D_3(n) = n^z \log n$ and $D_2(n) = F_{\max} n^z \log n$, for any set of hypercubes $\boldsymbol{C}$ that has been active and contained $\boldsymbol{x}_i(n')$ for some patients $n' \in \{1, \ldots, N\}$ of clinic i, the regret due to a near optimal clinic choosing a suboptimal expert for these patients when called by clinic i is bounded by $4(M-1)F_{\max}\beta_2$, where $\beta_2 = \sum_{t=1}^{\infty} 1/n^2$.

Proof: See Appendix A-D. ■

The next lemma bounds the regret due to clinic i choosing near optimal diagnostic actions for its patients up to the Nth patient.

Lemma 4: **Regret due to near-optimal experts.** Let $\mathcal{L}_{\boldsymbol{C},B}^i$, $B = 12/(L2^{-\alpha}) + 2$ denote the set of suboptimal actions for set of hypercubes $\boldsymbol{C}$. When LEX is run with parameters $p > 0$, $2\alpha/p \leq z < 1$, $D_1(n) = D_3(n) = n^z \log n$ and $D_2(n) = F_{\max} n^z \log n$, the regret due to near optimal diagnostic action selections in $\mathcal{L}_{\boldsymbol{C}_i(n),B}^i$ for patients $1 \leq n \leq N$ of clinic i in exploitation phases is bounded above by $O\left(N^{\frac{1+p-\alpha}{1+p}}\right)$.

Proof: See Appendix A-E. ■

Next, we combine the results from Lemmas 1, 2, 3 and 4 to obtain the regret bound for LEX.

Theorem 1: **Convergence rate to the optimal expert.** Let $\mathcal{L}_{\boldsymbol{C},B}^i$, $B = 12/(L2^{-\alpha}) + 2$ denote the set of suboptimal actions for set of hypercubes $\boldsymbol{C}$. When LEX is run with parameters $p = \frac{3\alpha + \sqrt{9\alpha^2 + 8\alpha}}{2}$, $z = 2\alpha/p < 1$, $D_1(n) = D_3(n) = n^z \log n$ and $D_2(n) = F_{\max} n^z \log n$, the regret of clinic i for its patients up to the Nth patient is upper bounded by $O\left(F_{\max} M N^{f_1(\alpha)} \log N\right)$, where $f_1(\alpha) = (2 + \alpha + \sqrt{9\alpha^2 + 8\alpha})/(2 + 3\alpha + \sqrt{9\alpha^2 + 8\alpha})$. Hence the average learning loss, i.e., $R_i(N)/N$ goes to zero.

Proof: For clinic i, 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 action selections 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 action selections 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. ■

From the result of Theorem 1, it is observed that the regret increases linearly with the number of clinics in the system and their number of experts (which $F_{\max}$ is an upper bound on). We note that the regret is the gap between the total expected reward of the optimal distributed policy that can be computed by a genie which knows the accuracy of every expert, and the total expected diagnostic reward of LEX. Since the performance of optimal distributed policy never gets worse as more clinics are added to the system or as more experts are introduced, the benchmark we compare our algorithm against with may improve. Therefore, the total reward of LEX may improve even if the regret increases with M, $|\mathcal{F}_i|$ and $F_{\max}$.

Theorem 1 gives a bound on the long-term performance of LEX. In a clinical setting, for interpreting the diagnosis recommendation provided by LEX, clinicians may want to know the confidence about the proposed diagnosis recommendation for the patient under consideration. LEX 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.

Corollary 1: **Confidence bounds on the diagnosis recommendation.** When LEX is run with parameters $p = \frac{3\alpha+\sqrt{9\alpha^2+8\alpha}}{2}$, $z = 2\alpha/p < 1$, $D_1(n) = D_3(n) = n^z \log n$ and $D_2(n) = F_{\max} n^z \log n$, we have the following confidence bounds on the diagnostic recommendation $\hat{y}_i(n)$: (i) If the prediction is made in an exploitation phase, then with probability

$$\min\left\{0, 1 - \frac{(|\mathcal{F}_i| + (M-1))(1+D)}{n^2} -2(M-1)F_{\max}\beta_2 n^{z/2-1}\right\},$$

the recommendation is made by a near optimal expert in set $\mathcal{K}_i - \mathcal{L}^i_{\boldsymbol{C},B}$, where $B = 12/(L2^{-\alpha})+2$ and $\beta_2 = \sum_{n=1}^{\infty} 1/n^2$. (ii) with the same probability as in (1), we have

$$\pi^d_{a_i(n)}(x^d) \geq \bar{r}^{i,d}_{a_i(n),C^d}(n) + c^i_{a_i(n)} - 3n^{-z/2} - L2^{l_{\max}(\boldsymbol{C})\alpha},$$

for all $d \in \mathcal{D}$, where $a_i(n)$ is the diagnostic action taken by clinic i for its nth patient and $\boldsymbol{C} = \boldsymbol{C}_i(n)$.

Proof: The proof is contained within the proof of Lemma 2. ■

Corollary 1 implies that when LEX exploits for a patient n, it can tell the clinician the probability that the chosen expert is one of the best (near-optimal) experts for the context of patient n using (1). Moreover, it can also tell the clinician a bound on the accuracy of the current diagnostic recommendation. This bound is given in (1), which says that for the patient population with type-d context x^d, the true accuracy of the best expert corresponding to the diagnostic action $a_i(n)$ will almost be as high as its estimated accuracy, i.e., $\bar{r}^{i,d}_{a_i(n),C^d}(n) + c^i_{a_i(n)}$. Using this information, the clinician can arrive at a decision: it can follow the recommendation of LEX, or it can find and assign another expert to the patient.

Remark 1: From Corollary 1 when $\alpha = 1$, the probability that a suboptimal expert is chosen for the nth patient when LEX exploits for the nth patient is $O(n^{-0.72})$. Although this goes to zero very quickly, the recommendation for the initial set of patients may not be very accurate. This is not a problem since prior knowledge can be incorporated to LEX. Assume that LEX is a priori trained with N_0 patients for each clinic. All the recommendations in this training set is done for the purpose of training LEX and does not affect a clinician's final decision on the patient. Then, for the nth patient that arrives to clinic i after the initial training, the probability that LEX chooses a suboptimal expert will be $O((n+N_0)^{-0.72})$, which can be made arbitrarily small by adjusting the initial training population.

V. Illustrative Results

In this section, we illustrate the functioning and performance of our algorithm by comparing it against several state-of-art online ensemble learning algorithms and other multi-armed bandit algorithms. While many of these algorithms are centralized and they cannot easily or at all be deployed in the envisioned distributed clinic setting, we compare against these various methods to highlight the merits of our proposed scheme, including the importance of using contextual (semantic) information in making decisions.

We consider a breast cancer data set provided by UCLA radiology department. The data set includes 45450 patients. 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 malignant. The score '4' is further divided into three subcategories with 4A indicating low suspicion of malignancy, 4B indicating intermediate suspicion and 4C indicating moderate concern. We focus only on the BI-RADS 4, 4A, 4B and 4C patients who need to be further monitored and/or screened to decide their cancer status (i.e. benign or malignant tumor). These patients are assigned to an expert, which decides whether or not to undergo a biopsy based on the patients' context, which includes their age, imaging modality, and breast density. Note that some instances of the context vector are missing for some patients. For instance, the breast density information is available for only 45% of the patients.

Unless stated otherwise, we assume that there are $M = 4$ clinics and $F_i = 2$ experts for each clinic $i \in \mathcal{M}$. A clinic can select one of its own experts without incurring any cost, while the cost of selecting another clinic is set to 0.01.

A. Algorithms that we compare against

-LinUCB [31] is a contextual bandit based algorithm which assumes that the expected reward of a diagnostic action is a linear combination of the components of the context vector. However, the coefficients in the linear combination is different for each diagnostic action and is unknown.

-Hybrid-ϵ (Hybrid) [32] combines the context (side) information with an ϵ-greedy algorithm, by extracting the history of context arrivals within a small region of the context space and running an ϵ-greedy algorithm within that space.

- Weighted Majority (WM) [33] is an offline algorithm that assigns and updates weights for the experts based on training data, and produces a final diagnosis recommendation by weighting the diagnosis recommendations of all the experts.

- Sliding Window Adaboost (AdaSliding) [34] is an online version of Adaboost [35], which aims to find the optimal weighting among the experts by an exponential weight update mechanism, where the weights are updated using a window of recent past observations and decisions.

All the algorithms above are centralized, i.e., they require a central clinic which has direct access to all the experts of all clinics. From the above algorithms, we modified LinUCB and Hybrid-ϵ such that they can run on the distributed setting we consider. For Weighted Majority and Sliding Window Adaboost we assume that there is a central clinic which has direct access to all the experts of all clinics.

B. Performance of LEX

While LinUCB and Hybrid-ϵ are bandit algorithms (i.e., they require a diagnosis only from the expert they select), AdaSliding and AdaWeighted require the diagnostic recommendations of all the experts in the system for each patient. Hence, these algorithms are run on a centralized system in which the clinic has access to all experts.

Results on the diagnostic accuracy: The Comparison between LEX and the above algorithms is given in Table II. As the performance metric, we use the diagnostic accuracy (i.e., the percentage of patients that are correctly diagnosed by LEX). LEX outperforms other learning algorithms by having more than 13% diagnostic accuracy than the best of the other methods. Moreover, as the number of patients N increases, the diagnostic accuracy increases because LEX learns the expertise of the experts with a higher accuracy as more patients arrive. The poor performance of LinUCB and Hybrid-ϵ algorithms is due to the fact that they don't have the training phase that LEX have, and they learn considering all the contexts in the context vector, rather than learning for the most relevant context as LEX does.

Results on cooperation among the clinics: In order to assess the effect of cooperation between clinics, we simulate the performance of LEX for different numbers of clinics that clinic i can forward its patient's context for diagnosis recommendation. As shown in Table III diagnosis accuracy of LEX increases with the number of clinics that clinic i is connected to. This is due to the diversity of the expertise among different clinics. While a clinic can be good at making diagnosis recommendation to patients with a specific type of context, another clinic may be better specialized for other types of contexts.

TABLE III
DIAGNOSTIC ACCURACY OF LEX AS A FUNCTION OF THE NUMBER OF CLINICS

Number of clinics	M=1	M=2	M=4
LEX	75.96%	81.09%	83.32%

Results on costs associated with cooperation: How and when the clinics cooperate with each other depends on several factors including the expertise of the clinics, contexts of the patients and costs of cooperation (delay, money, etc.). Here, we evaluate the percentage of times a clinic cooperates with the other 4 clinics in exploitations using LEX, as a function of the cost of choosing another clinic for diagnosis recommendation. We assume this cost is the same for all clinics and denote it by c. As seen in Table IV, the percentage of cooperation decreases as the cost increases, and reaches zero when the cost exceeds some threshold. When the cost is too high, asking for expertise of another clinic is not advantageous, even when it improves the diagnostic accuracy.

TABLE IV
COOPERATION % VS. COOPERATION COST. COOPERATION % IS THE PERCENTAGE OF TIMES ANOTHER CLINIC IS CALLED IN EXPLOITATIONS.

Cost c	0.01	0.05	0.1	0.2	0.5
Cooperation	12.67%	7.68%	5.53%	0%	0%

Results on delayed health state observations: For most cases, it may not be possible observe the true health state of the patient just after the diagnostic decision is made. When the decision is "malignant", usually a biopsy is performed in a short amount of time, and this reveals the true health state of the patient. However, when the decision is "benign", the patient usually waits until the next screening, without any immediate action.

We simulate this by introducing a delay d_n, in terms of the number of patients that have arrived after the nth patient before the true health state of the nth patient is revealed. We assume that $d_n \sim \text{geometric}(\lambda)$, for some parameter $\lambda > 0$. In Table V, the performance of LEX as a function of the delay is shown in terms of parameter λ. Smaller values of λ imply larger delay, but as seen from the table, the performance of LEX is only slightly affected by the delay.

TABLE V
TRADEOFF BETWEEN DIAGNOSTIC ACCURACY AND DELAY

λ	0.02	0.01	0.004	0.002
LEX	81.48%	80.43%	80.22%	79.18%

VI. CONCLUSION

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, which means that the best diagnosis mechanism (whether a human expert or a CDSS) for each context is perfectly learned. Moreover, the proposed algorithm LEX learns the best expert for treating a patient with a specific context not only within a clinic but also across all clinics; hence, its performance is better than the performance of the best expert within any given clinic. In a clinical deployment of the proposed system, diagnosis recommendations made by the system will be examined by a clinician before the final prediction is made. This will provide an additional layer of safety. For any patient, it is the clinician's discretion whether to rely on or to disregard the recommendation of the LEX algorithm. Future work includes designing algorithms that can track the changes in a clinician's performance by exploiting a recent time window of patient histories.

APPENDIX A
PROOF OF THE LEMMAS

A. 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 5: **A bound on the level of active hypercubes.** All the active hypercubes $\mathcal{A}^d_i(n)$ 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 $n/A > 1$, $A \frac{2^{p(l+1)}-1}{2^p-1} < n \Rightarrow 2^{pl} < \frac{n}{A} \Rightarrow l < \frac{\log_2 n}{p}$. ■

B. Proof of Lemma 1

This directly follows from the number of trainings and explorations that are required before any diagnostic action can be exploited (see definition of $S^i_{C_i(n)}(n)$). If the diagnosis recommendation at any training or exploration phase is incorrect or a high cost diagnostic action is chosen, clinic i 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 cost of a diagnostic action can at most be one.

TABLE II
COMPARISION OF LEX WITH STATE-OF-THE-ART LEARNING ALGORITHMS IN TERMS OF DIAGNOSTIC ACCURACY

Patients	Type	Type	Type	$N = 1000$	$N = 3000$	$N = 5439$
LEX	Distributed	Contextual	Online	80.03%	82.49%	83.32%
LinUCB	Distributed	Contextual	Online	63.03%	65.93%	66.43%
Hybrid	Distributed	Contextual	Online	63.15%	65.53%	67.91%
AdaSliding	Centralized	-	Online	66.10%	71.20%	73.16%
WM	Centralized	-	Offline	60.50%	59.93%	59.48%

C. Proof of Lemma 2

Let Ω denote the space of all possible outcomes, and w be a sample path. The event that the LEX exploits when $\boldsymbol{x}_i(n) \in \boldsymbol{C}$ is given by $\mathcal{W}^i_{\boldsymbol{C}}(n) := \{w : S^i_{\boldsymbol{C}}(n) = \emptyset, \boldsymbol{x}_i(n) \in \boldsymbol{C}, \boldsymbol{C} \in \mathcal{A}_i(n)\}$. We will bound the probability that LEX chooses a suboptimal action for clinic i in an exploitation phase when i's 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 i for its patients in exploitation steps using LEX. Recall that reward loss in every step in which a suboptimal action is chosen can be at most 2.

Let $\mathcal{V}^i_{k,\boldsymbol{C}}(n)$ be the event that a suboptimal action k is chosen for the set of hypercubes $\boldsymbol{C}$ by clinic i for its nth patient. For $k \in \mathcal{K}_i \cap \mathcal{F}_i$, let $\mathcal{E}^i_{k,\boldsymbol{C}}(n)$ be the set of rewards observed by clinic i from action k for its own patients whose context vectors is in the active set $\boldsymbol{C}$ up to the nth patient. For $j \in \mathcal{K}_i \cap \mathcal{M}_{-i}$, let $\mathcal{E}^i_{j,\boldsymbol{C}}(n)$ be the set of rewards observed from selections of clinic j for patients $n' \in \{1, \ldots, n\}$ of clinic i with context vectors in the active set $\boldsymbol{C}$ such that $T^i_{1,j,l}(n') > D_2(n')$. Let $\mathcal{B}^i_{j,\boldsymbol{C}}(n)$ be the event that at most n^{ϕ} observations in $\mathcal{E}^i_{j,\boldsymbol{C}}(n)$ are collected from suboptimal actions of clinic j. For $k \in \mathcal{K}_i \cap \mathcal{F}_i$ let $\mathcal{B}^i_{k,\boldsymbol{C}}(n) := \Omega$. We generate two different artificial i.i.d. processes to bound the probabilities related to deviation of sample mean reward estimates $\bar{r}^{i,d}_{k,C^d}(n)$, $k \in \mathcal{K}_i$, $d \in \mathcal{D}$ 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 $\overline{\mu}^d_{k,C^d}$, the other one is the *worst* process in which the rewards are generated according to a bounded i.i.d. process with expected reward $\underline{\mu}^d_{k,C^d}$. Let $\bar{r}^{b,i,d}_{k,C^d}(n)$ denote the sample mean of the n samples from the best process and $\bar{r}^{w,i,d}_{k,C^d}(n)$ denote the sample mean of the n samples from the worst process. We have for any $k \in \mathcal{L}^i_{\boldsymbol{C},B}$

$$
\begin{aligned}
&P\left(\mathcal{V}^i_{k,\boldsymbol{C}}(n), \mathcal{W}^i_{\boldsymbol{C}}(n)\right) \\
&\leq P\left(\max_{d\in\mathcal{D}} \bar{r}^{b,i,d}_{k,C^d}(T^{i,d}_{k,C^d}(n)) \geq \overline{\mu}_{k,\boldsymbol{C}} + H_n, \mathcal{W}^i_{\boldsymbol{C}}(n)\right) \\
&+ P\Bigg(\max_{d\in\mathcal{D}} \bar{r}^{b,i,d}_{k,C^d}(T^{i,d}_{k,C^d}(n)) \geq \bar{r}^{w,i,d^*(\boldsymbol{C})}_{k^*(\boldsymbol{C}),C^{d^*(\boldsymbol{C})}}(T^{i,d^*(\boldsymbol{C})}_{k^*(\boldsymbol{C}),C^{d^*(\boldsymbol{C})}}(n)) \\
&-2n^{\phi-1}, \max_{d\in\mathcal{D}} \bar{r}^{b,i,d}_{k,C^d}(T^{i,d}_{k,C^d}(n)) < \overline{\mu}_{k,\boldsymbol{C}} + L2^{-l_{\max}(\boldsymbol{C})\alpha} \\
&+H_n + 2n^{\phi-1}, \bar{r}^{w,i,d^*(\boldsymbol{C})}_{k^*(\boldsymbol{C}),C^{d^*(\boldsymbol{C})}}(T^{i,d^*(\boldsymbol{C})}_{k^*(\boldsymbol{C}),C^{d^*(\boldsymbol{C})}}(n)) \\
&> \underline{\mu}_{k^*(\boldsymbol{C}),\boldsymbol{C}} - L2^{-l_{\max}(\boldsymbol{C})\alpha} - H_n, \mathcal{W}^i_{\boldsymbol{C}}(n)\Bigg) \\
&\quad + P\left(\bar{r}^{w,i,d^*(\boldsymbol{C})}_{k^*(\boldsymbol{C}),C^{d^*(\boldsymbol{C})}}(T^{i,d^*(\boldsymbol{C})}_{k^*(\boldsymbol{C}),C^{d^*(\boldsymbol{C})}}(n)) \leq \underline{\mu}_{k^*(\boldsymbol{C}),\boldsymbol{C}} - H_n \right. \\
&\quad \left. +2n^{\phi-1}, \mathcal{W}^i_{\boldsymbol{C}}(n)\right) + P((\mathcal{B}^i_{k,\boldsymbol{C}}(n))^c),
\end{aligned}
$$

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

$$4n^{\phi-1} + 2H_n \leq (B-2)L2^{-l_{\max}(\boldsymbol{C})\alpha}.$$

By Lemma 5, (A-C) holds when

$$4n^{\phi-1} + 2H_n \leq (B-2)L2^{-\alpha}n^{-\alpha/p}.$$

For $H_n = 4n^{\phi-1}$, $\phi = 1 - z/2$, $z \geq 2\alpha/p$ and $B = 12/(L2^{-\alpha}) + 2$, (A-C) holds by which (A-C) is equal to zero. Also by using a Chernoff-Hoeffding bound we can show that

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

and

$$
\begin{aligned}
&+ P\left(\bar{r}^{w,i,d^*(\boldsymbol{C})}_{k^*(\boldsymbol{C}),C^{d^*(\boldsymbol{C})}}(T^{i,d^*(\boldsymbol{C})}_{k^*(\boldsymbol{C}),C^{d^*(\boldsymbol{C})}}(n)) \leq \underline{\mu}_{k^*(\boldsymbol{C}),\boldsymbol{C}} - H_n \right. \\
&\left. +2n^{\phi-1}, \mathcal{W}^i_{\boldsymbol{C}}(n)\right) \leq 1/n^2.
\end{aligned}
$$

We also have $P(\mathcal{B}^i_{k,\boldsymbol{C}}(n)^c) = 0$ for $k \in \mathcal{F}_i$ and $P(\mathcal{B}^i_{j,\boldsymbol{C}}(n)^c) \leq E[X^i_{j,\boldsymbol{C}}(n)]/n^{\phi} \leq 2F_{\max}\beta_2 n^{z/2-1}$. for $j \in \mathcal{M}_{-i}$, where $X^i_{j,\boldsymbol{C}}(n)$ is the number of times a suboptimal expert of clinic j is selected when clinic i calls clinic j in exploration and exploitation phases for patients of clinic i with context vectors in the set of hypercubes $\boldsymbol{C}$ that are active for the nth patient. Combining all of these we get $P\left(\mathcal{V}^i_{k,\boldsymbol{C}}(n), \mathcal{W}^i_{\boldsymbol{C}}(n)\right) \leq (1+D)/n^2$, for $k \in \mathcal{F}_i$ and $P\left(\mathcal{V}^i_{j,\boldsymbol{C}}(n), \mathcal{W}^i_{\boldsymbol{C}}(n)\right) \leq (1+D)/n^2 + 2F_{\max}\beta_2 n^{z/2-1}$, for $j \in \mathcal{M}_{-i}$. We get the final bound by summing these probabilities from $n = 1$ to N.

D. Proof of Lemma 3

Let $X^i_{j,\boldsymbol{C}}(N)$ denote the random variable which is the number of times a suboptimal expert of clinic $j \in \mathcal{M}_{-i}$ is chosen in exploitation phases of clinic i when $\boldsymbol{x}_i(n')$ is in set $\boldsymbol{C} \in \mathcal{A}_i(n')$ for $n' \in \{1, \ldots, N\}$. It can be shown that $E[X^i_{j,\boldsymbol{C}}(N)] \leq 2F_{\max}\beta_2$. Thus, the contribution to the regret from suboptimal actions of clinic j is bounded by $4F_{\max}\beta_2$. We get the final result by considering the regret from all $M-1$ other clinics.

E. Proof of Lemma 4

The following lemma bounds the per-patient (one-step) regret to clinic i from choosing near optimal actions. This lemma is used later to bound the total regret from near optimal actions.

Lemma 6: **One-step regret due to near-optimal actions for a set of hypercubes.** Let $\mathcal{L}^i_{\boldsymbol{C},B}$, $B = 12/(L2^{-\alpha}) + 2$ denote the set of suboptimal actions for set of hypercubes $\boldsymbol{C}$. When LEX is run with parameters $p > 0$, $2\alpha/p \leq z < 1$, $D_1(n) = D_3(n) = n^z \log n$ and $D_2(n) = F_{\max} n^z \log n$, for any set of hypercubes $\boldsymbol{C}$, the one-step regret of clinic i from choosing one of its near optimal actions is bounded above by $BL2^{-l_{\max}(\boldsymbol{C})\alpha}$, while the one-step regret of clinic i from choosing a near optimal clinic which chooses one of its near optimal experts is bounded above by $2BL2^{-l_{\max}(\boldsymbol{C})\alpha}$.

Proof: For the nth patient of clinic i if $\boldsymbol{x}_i(n) \in \boldsymbol{C} \in \mathcal{A}_i(n)$, the per-patient regret of any near optimal expert of any near optimal clinic $j \in \mathcal{M}_{-i}$ is bounded by $2BL2^{-l_{\max}(\boldsymbol{C})\alpha}$ by the definition of $\mathcal{L}^i_{\boldsymbol{C},B}$. Similarly, the per-patient regret of any near optimal expert $k \in \mathcal{F}_i$ is bounded by $BL2^{-l_{\max}(\boldsymbol{C})\alpha}$. ■

For the nth patient of clinic i, for the set of active hypercubes $\boldsymbol{C}_i(n)$ that the patient's context vector belongs to, $l_{\max}(\boldsymbol{C}_i(n))$ is at least the level of the active hypercube $x^d_i(n) \in C^d_i(n)$ for some type-d context. Since a near optimal action's regret for the nth patient is upper bounded by $2BL2^{-l_{\max}(\boldsymbol{C}_i(n))\alpha}$, the total regret due to near optimal actions up to the Nth patient is upper bounded by $2BL\sum_{n=1}^{N} 2^{-l_{\max}(\boldsymbol{C}_i(n))\alpha} \leq 2BL\sum_{n=1}^{N} 2^{-l(C^d_i(n))\alpha}$. Let $l_{\max,u}$ 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 $l_{\max,u} < 1+\log_2(N)/(1+p)$. $\sum_{n=1}^{N} 2^{-l(C^d_i(n))\alpha}$ takes its greatest value when type-d context up to the Nth patient is uniformly distributed in $\mathcal{X}_d$. Therefore we have $\sum_{n=1}^{N} 2^{-l(C^d_i(n))\alpha} \leq \frac{A2^{2(1+p-\alpha)}}{2^{1+p-\alpha}-1} N^{\frac{1+p-\alpha}{1+p}}$.

References

[1] P. Sheu, H. Yu, C. Ramamoorthy, A. K. Joshi, and L. A. Zadeh, *Semantic Computing*. John Wiley & Sons, 2011.

[2] C. C. Wang, D. A. Hecht, P. C.-Y. Sheu, and J. J. Tsai, "Semantic computing and drug discovery - a preliminary report," *2012 IEEE Sixth International Conference on Semantic Computing*, vol. 0, pp. 453–458, 2013.

[3] C. C. Bennett and K. 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.

[4] J. G. Elmore, D. L. Miglioretti, L. M. Reisch, M. B. Barton, W. Kreuter, C. L. Christiansen, and S. 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.

[5] M. N. Gurcan, B. Sahiner, N. Petrick, H.-P. Chan, E. A. Kazerooni, P. N. Cascade, and L. 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.

[6] J. Tang, R. M. Rangayyan, J. Xu, I. El Naqa, and Y. 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.

[7] K. Ganesan, U. Acharya, C. K. Chua, L. C. Min, K. Abraham, and K. Ng, "Computer-aided breast cancer detection using mammograms: a review," *Biomedical Engineering, IEEE Reviews in*, vol. 6, pp. 77–98, 2013.

[8] M. F. Ganji and M. S. Abadeh, "A fuzzy classification system based on ant colony optimization for diabetes disease diagnosis," *Expert Systems with Applications*, vol. 38, no. 12, pp. 14 650–14 659, 2011.

[9] A. X. Garg, N. K. Adhikari, H. McDonald, M. P. Rosas-Arellano, P. Devereaux, J. Beyene, J. Sam, and R. B. 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.

[10] A. D. Black, J. Car, C. Pagliari, C. Anandan, K. Cresswell, T. Bokun, B. McKinstry, R. Procter, A. Majeed, and A. Sheikh, "The impact of eHealth on the quality and safety of health care: a systematic overview," *PLoS Medicine*, vol. 8, no. 1, p. e1000387, 2011.

[11] W. Hsu, R. K. Taira, S. El-Saden, H. Kangarloo, and A. A. 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.

[12] K. Kerlikowske, P. A. Carney, B. Geller, M. T. Mandelson, S. H. Taplin, K. Malvin, V. Ernster, N. Urban, G. Cutter, R. Rosenberg *et al.*, "Performance of screening mammography among women with and without a first-degree relative with breast cancer," *Annals of Internal Medicine*, vol. 133, no. 11, pp. 855–863, 2000.

[13] H. van der Linden, D. Kalra, A. Hasman, and J. Talmon, "Inter-organizational future proof ehr systems: a review of the security and privacy related issues," *International Journal of Medical Informatics*, vol. 78, no. 3, pp. 141–160, 2009.

[14] T. Gu, H. K. Pung, and D. Q. Zhang, "A middleware for building context-aware mobile services," in *Vehicular Technology Conference, 2004. VTC 2004-Spring. 2004 IEEE 59th*, vol. 5. IEEE, 2004, pp. 2656–2660.

[15] X. Wang, J. S. Dong, C. Chin, S. R. Hettiarachchi, and D. Zhang, "Semantic space: An infrastructure for smart spaces," *Computing*, vol. 1, no. 2, pp. 67–74, 2002.

[16] M. Sewell, "Ensemble learning," *RN*, vol. 11, no. 02, 2008.

[17] E. Alpaydin, *Introduction to machine learning*. The MIT Press, 2004.

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

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

[20] A. Lazarevic and Z. 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.

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

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

[23] T. Lu, D. Pál, and M. Pál, "Contextual multi-armed bandits," in *International Conference on Artificial Intelligence and Statistics*, 2010, pp. 485–492.

[24] W. Chu, L. Li, L. Reyzin, and R. E. Schapire, "Contextual bandits with linear payoff functions," in *Proc. of the International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2011.

[25] B. Chen, R. Jiang, T. Kasetkasem, and P. K. Varshney, "Channel aware decision fusion in wireless sensor networks," *Signal Processing, IEEE Transactions on*, vol. 52, no. 12, pp. 3454–3458, 2004.

[26] G. Mateos, J. A. Bazerque, and G. B. Giannakis, "Distributed sparse linear regression," *Signal Processing, IEEE Transactions on*, vol. 58, no. 10, pp. 5262–5276, 2010.

[27] H. Kargupta, B. Park, D. Hershberger, and E. Johnson, "Collective data mining: A new perspective toward distributed data mining," *Advances in Distributed and Parallel Knowledge Discovery*, no. part II, pp. 131–174, 1999.

[28] H. Zheng, S. R. Kulkarni, and H. Poor, "Attribute-distributed learning: models, limits, and algorithms," *Signal Processing, IEEE Transactions on*, vol. 59, no. 1, pp. 386–398, 2011.

[29] T. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," *Advances in Applied Mathematics*, vol. 6, pp. 4–22, 1985.

[30] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine Learning*, vol. 47, p. 235256, 2002.

[31] L. Li, W. Chu, J. Langford, and R. E. Schapire, "A contextual-bandit approach to personalized news article recommendation," in *Proc. of the 19th international conference on World wide web*. ACM, 2010, pp. 661–670.

[32] D. Bouneffouf, A. Bouzeghoub, and A. L. Gançarski, "Hybrid-ε-greedy for mobile context-aware recommender system," in *Advances in Knowledge Discovery and Data Mining*. Springer, 2012, pp. 468–479.

[33] N. Littlestone and M. K. Warmuth, "The weighted majority algorithm," in *Foundations of Computer Science, 1989., 30th Annual Symposium on*. IEEE, 1989, pp. 256–261.

[34] H. Wang, W. Fan, P. S. Yu, and J. Han, "Mining concept-drifting data streams using ensemble classifiers," in *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2003, pp. 226–235.

[35] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," in *Computational learning theory*. Springer, 1995, pp. 23–37.