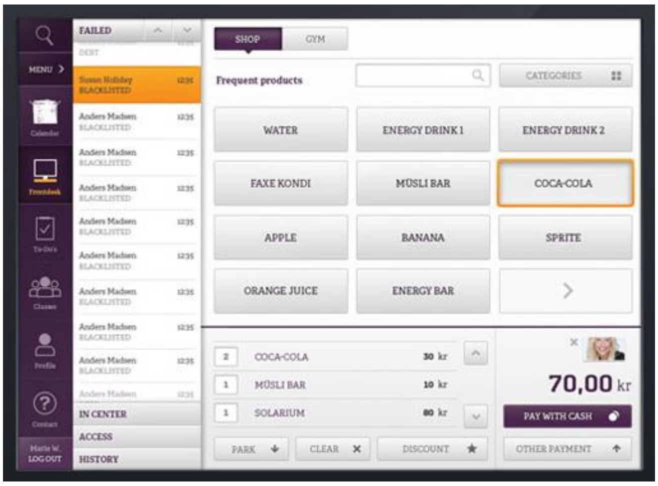
Industrial PhD project description

Basic information

Project title	ONFIT: Online learning for fitness centers
Company	Exerp
University, center/department	University of Copenhagen, Department of Computer Science (DIKU)
Industrial PhD candidate	–
Any third party eligible for subsidy	–
Any third party not eligible for subsidy	–

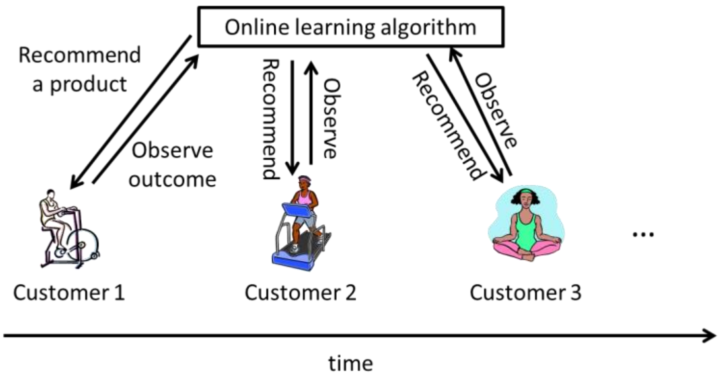
A. State-of-the-art and theoretical background

General Background



Exerp specializes in providing membership management software for fitness centers. Its main product SAAS is used by 839 fitness clubs in seven countries in Europe that jointly serve over two million customers. There is enormous amount of information processed by SAAS, from customers' subscriptions to fitness class enrolments to drinks and snacks purchases, and many other details involved in fitness club operation (a snapshot from SAAS is presented on the right). The goal of this project is to exploit this information to provide more revenue for fitness clubs and better value for their customers. We plan to achieve this goal by designing a *personalized recommender system*. The recommender system will improve the efficiency of fitness clubs by *improving the match between the products and the customers*. It will be achieved by *recommending the right products to the right customers*.

The intelligent core of the proposed recommender system will be based on state-of-the-art *online learning* algorithms that will be developed in this project. Online learning is a subfield of machine learning modeling situations in which learning and information acquisition occur simultaneously. In the case of fitness industry we have a continuous flow of transactions (see illustration on the right) and the goal is to achieve continuous improvement of recommendations with time. Online learning has already made significant impact in online advertising [CL12][1], online news recommendation [LCLS10], online retail (in online stores), and many other domains, but it has not yet been applied in the fitness industry.



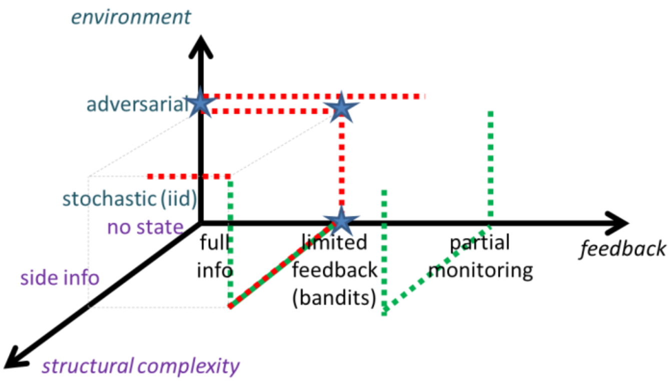
State-of-the-art and Theoretical Background

We identify three major characteristics of online learning problems (see figure on the next page):

---

[1] All bibliographic references are provided at the end of this document.

1. The first is the amount of information observed after every action (the *feedback* axis in the figure). It can be full information, limited (a.k.a. bandit) feedback, or partial feedback (partial monitoring). An example of full information feedback is the stock market, where we observe the rate of all stocks and in retrospect can evaluate the performance of any investment strategy. An example of bandit feedback is medical treatment – when patient receives a particular treatment we observe the response to that treatment, but we do not know what would have happened had we chosen a different treatment. This lack of information makes direct retrospective evaluation of alternative strategies impossible. More rigorously, in bandit feedback we observe the outcome of the action that was performed, but not of any alternative action. Partial feedback is a more general feedback model, where the observation is related to the quality of the performed action in an arbitrary way. For example, in dynamic pricing we only observe whether the price was acceptable for the customer, but we do not observe the maximal price the customer had in mind.



2. The second characteristic is the adversariality of the environment (the *environment* axis). We distinguish between i.i.d. (a.k.a. "stochastic"[2]) and adversarial environments. For example, a series of patients is an instance of an i.i.d. environment, whereas spam filtering is an adversarial online learning problem.

3. The third characteristic is the structural complexity of a problem (the *structural complexity* axis). In the simplest case the problem is stateless. A harder case is when there is side information (state) that has to be taken into account. For example, the medical record of a patient represents its state.

Most of existing results in online learning start with a set of assumptions about the feedback model, environment adversariality, and structural complexity and solve the problem under these assumptions [CBL06, BCB12]. For example, prediction with expert advice, adversarial multiarmed bandits, and stochastic multiarmed bandits correspond to the three blue stars on the figure (in clockwise direction order).

B. Research objective

The scientific goal of this project is to advance the theory of online learning for recommender systems and to increase their effectiveness in the application scenarios occurring in the fitness industry. The problem with existing theoretical results is that practical problems rarely match idealistic theoretical assumptions. If we think about the fitness industry, it is not really adversarial, because customers mainly follow their habits, much like patients in medical treatments or viewers in movie recommendation domains. But it is also not purely stochastic, because there are seasonal variations, trends, and also truly adversarial events, such as marketing campaigns of competing fitness clubs. Therefore, treating fitness industry with stochastic models is risky, but treating it with adversarial models does not allow exploiting the simplicity of the problem.

In the fitness industry the feedback is also not purely bandit, because we can recommend more than a single product and get more information. And the side information is usually only partially relevant, so the location along the structural complexity axis can also vary. Thus, existing theoretical models study isolated points in the space of online learning problems, whereas practical problems we are facing fall between these isolated points.

In a number of recent works, we designed algorithms that interpolate between points in the space of online learning problems (shown by red dashed lines in the figure). In particular, we designed "one practical algorithm for both stochastic and adversarial bandits" [SS14]; an interpolation between stateless bandits and bandits with side information in the stochastic regime [SAL+11]; and two models that interpolate between bandit and full information feedback in adversarial stateless problems ("prediction with limited advice" and "multiarmed bandits with paid observations") [SBCA14]. The latter was further extended by Kale to bandits with limited advice [Kal14]. The practical advantage of such interpolations is that they allow to relax the

[2] In the field of online learning the word "stochastic" is frequently used as a synonym to "i.i.d." and it should not be confused with "stochastic processes" if the reader is coming from a different field.

assumptions and to exploit the simplicity of a problem whenever it is present without compromising on the worst-case performance guarantees.

The objective of the proposal is to continue this line of research and design new interpolations that are relevant for personalization problems in the fitness industry. In particular, we plan to improve the interpolation between stochastic bandits with side information and stateless bandits; to derive an interpolation between stochastic and adversarial bandits with side information; and to derive interpolations between stochastic and adversarial partial monitoring games and between partial monitoring games with and witout side information (see green dashed lines on the figure). These algorithms will power the recommender systems we will add to Excerp's IT platforms for large-scale fitness operators. The relevance of the proposed models for the fitness industry is described in more details in Section D. of the proposal.

C. Commercial perspectives

The fitness industry has become big business. For example, the number of customers in Denmark has steadily been growing throughout the last 20 years (Exerp's internal data), but more importantly, the market leaders have changed from single club owners into holding companies and insurance companies who have all seen the value of exercise in the area of obesity. Examples of consolidation and change in ownership include CCMP Capital investing in PureGym in 2013, Tryghedsgruppen investing in SATS (Health and fitness nordic) in 2012, and Parken Sport and Entertainment investing in Fitnessdk. Fitnessworld.dk, which has all along been privately owned, has over the years purchased private individual clubs and small chains such as Fitness One and Enjoy Fitness in 2012, and is now the largest Danish fitness chain with more then 110 clubs. The fitness industry has matured from underdog into an important player in the world market economy with a yearly European revenue in 2013 of almost US$32.9 billion based on 44 million members belonging to more than 48,000 health clubs in Europe (Melissa Rodriguez, IHRSA's senior research manager, from the newly published 2014 IHRSA Global Report). Exerp is the membership-software supplier to all of the above mentioned fitness chains.

Personalization makes the connection between the products and the customers more efficient. It allows identification of consumer niches, development of niche products, and their sales. The benefits of personalization to the industry will be tremendous:

- The customers will be offered the products they really want and will be satisfied.
- As a result, the number of customers and the number of purchases by each customer will grow and thus the revenue will grow.
- Personalization will help in design and sales of new products.

Exerp's IT systems currently have no online learning functionality, even though Exerp is the leading software supplier for the fitness market in Europe. Exerp has no doubt that this unexplored research field will have a major impact on the future of the industry (Co-founder Jacob Wiberg, Exerp).

This is the first collaboration of Exerp with a research institution and the first attempt to apply rigorous statistical analysis to the wealth of data that Exerp holds. Therefore, it is also expected that the project will contribute to the profile of the company and promote it as a modern data-oriented enterprise. The envisioned recommender systems, which increase the revenue of the fitness centers, will be a decisive competitive advantage of Exerp's products. They will further strengthen the leadership of Exerp, both in Denmark and abroad, and provide opportunities for expansion to new markets.

D. Project description

The goal of this project is to introduce personalization into the fitness industry. In short, personalization aims at *offering the right products to the right customers*. Successful personalization strategies form the core of commercial success of companies such as Amazon, Netflix, and Google. One of the fundamental trade-offs in the fitness industry, as in many other industries, is the trade-off between choice offered to the clients and their ability to find what they really need. On the one side, having large choice is good, because every customer can potentially find the right product for his or her needs. On the other side, having large choice is bad, because in makes the search process harder. There is also a limit on the number of different products that can be offered, depending on the number of potential customers. Fast and successful matching between products and customers is the key for improving efficiency of the fitness industry.

Fitness industry offers a wide range of products. In this project we will focus on personalized recommendation of group fitness classes and on personalized recommendation of over-the-counter products (like snack bars and energy drinks). The two categories of products have very different nature and provide a good opportunity to consider personalization from different angles. We start with describing our planned approach to personalization of group fitness classes and then continue to over-the-counter products.

Personalized recommendation of group fitness classes (Work Package 1 – WP1)

Group fitness classes are characterized by the following parameters:

- The type of the class (yoga, spinning, boxing, etc.).
- The instructor of the class. We consider this to be more important feature than the type of the class in the sense that good instructors attract more participants irrespective of the type of the class. In other words, people can chose a class because of a good instructor rather than their prior interest in particular sport. It is a little bit like a film director name vs. the film genre. We also note that the match between instructor and a class participant is a personal match. Not all instructors are universally good. Some instructors are good for older people, some are good for younger people, some are good for beginners, and some are good for advanced participants, and so on. The role of personalized recommendation system is to offer the right instructor.
- The participants of the class. From our experience with movie recommendation systems [ST10, Netflix] we know that the identity of participants of a class is the best description of a class. It is the most valuable source of information for personalized recommendations. Assume that a person X is not enrolled on class Y, but X takes many other classes that are taken by the participants of Y. Then we can safely recommend Y to X.

The problem of personalized recommendation of group fitness classes fits well into the multiarmed bandits with side information model [SAL+11]. The arms of the bandit are the different fitness classes that can be offered to the customer and the side information are all the parameters mentioned above. The side information is abundant and most likely only little part of it is actually relevant for prediction. We will approach the personalization problem through the following steps:

Step 1: The algorithm from [SAL+11] allows to filter the relevant side information and its performance depends on the amount of relevant side information rather than the total amount of side information. We will start with applying this algorithm to the problem.

Step 2: We will compare the results of the algorithm of [SAL+11] to alternative algorithms for the problem that are based on graph clustering [CBGZ13, GLZ14]. Specifically, the approach of [GLZ14] assumes that there are clusters of customers with identical preferences and tries to identify the clusters.

Step 3: The approach of [GLZ14] is a bit more rigid compared to [SAL+11], since it assumes hard association of customers with clusters, whereas in [SAL+11] there is no hard partition of the customers. [GLZ14] also assume that the environment is adversarial, whereas [SAL+11] assume that the environment is stochastic. However, the performance scaling of the algorithm of [SAL+11] with the number of prediction rounds is suboptimal, because of problems with variance control. Since 2011 when the algorithm was proposed a number of alternative strategies for variance control and ways of avoiding high-variance problems were developed [Bub10, BCB12, SS14]. The third step of the project will be to adapt these new strategies to improve the algorithm of [SAL+11]. This will provide even better interpolation between bandits with side information and stateless bandits.

Step 4: The fourth step will be to combine the improved algorithm of [SAL+11] with the exploration technique of [SS14] in order to derive an algorithm capable of operating in both stochastic and adversarial regimes.

The final result will be a personalized recommendation algorithm that is capable of filtering the relevant side information and robust against deviations from the i.i.d. model. It will exploit the simplicity of the data whenever present without compromising on the worst-case performance.

Personalized recommendation of over-the-counter products (Work Package 2 – WP2)

The main control lever used by fitness centers to control purchases of drinks and snacks are personal discounts. Essentially, it means that we are talking about dynamic pricing model – we are looking for the minimal discount that will be sufficient for securing the transaction. Incorporation of information about the customer into the problem should allow improving the performance even more. Exerp already has a platform that can identify when customers are training and send targeted discounts to their smartphones. However, the infrastructure lacks intelligent algorithm for picking the discounts. The algorithm will be developed through the following steps:

Step 1: Dynamic pricing is a special case of a partial monitoring game [CBL06]. The project will start with implementing the simplest algorithm for adversarial partial monitoring games [CBL06].

Step 2: The second step will be to implement the algorithm for stochastic partial monitoring games [BZS12] and to compare it with the algorithm for adversarial partial monitoring games.

Step 3: The third step will be to extend the ideas on interpolation between adversarial and stochastic bandits [SS14] to partial monitoring games.

Step 4: This step will study the usage of side information in dynamic pricing. We will start with implementing the algorithm of [BS12].

Step 5: We will continue with extending the ideas on adaptive usage of side information from bandits [SAL+11] to stochastic partial monitoring games.

Step 6: Finally, we will work on robust algorithms for partial monitoring games with side information. Namely, we will design algorithms that are able to exploit the simplicity of dynamic pricing with side information in stochastic environments without compromising on the performance if the stochastic assumption is violated.

Data collection and other technical details (Work Package 0 – WP0)

As already mentioned, when feedback is limited or partial it is impossible to evaluate quality of alternative strategies directly. In particular, it is impossible to evaluate the quality of strategy A when the data was collected using strategy B. This makes offline comparison of different strategies tricky. Fortunately, there is an easy fix for this problem, known as *importance-weighted sampling*. Let us briefly explain the idea behind importance-weighted sampling. Imagine we have K actions. When we play action $i \in \{1, \dots, K\}$ we observe the quality of that action, denote it by $f(i)$, but not of any other action. The idea behind importance-weighted sampling is to pick the action to play at random according to a distribution $p(i)$ and to substitute $f(i)$ with an estimate

$$g(i) = \begin{cases} \frac{f(i)}{p(i)}, & \text{if action } i \text{ was played.} \\ 0, & \text{otherwise.} \end{cases}$$

Note that (as long as $p(i)$ is non-zero) $g(i)$ is an unbiased estimate of $f(i)$, because $\mathbb{E}[g(i)] = p(i)\frac{f(i)}{p(i)} = f(i)$. Thus, importance-weighted sampling allows, to a certain degree, to circumvent the problem of offline evaluation.

The project will start with the following steps that will improve the process of data collection at Exerp and prepare ground for presenting recommendations / discounts to the customers. This part of the project will be done in the first year of PhD with support from Laurent Czinczenheim, the COO of Exerp.

Step 1: Implement a basic interface for presenting recommendations / discounts to customers.

Step 2: Implement the basic ε-greedy recommendations / discounts strategy. The ε-greedy strategy operates in the following way. With probability (1- ε) it suggests the default (or most popular) option and with probability ε it suggests any other available option, uniformly at random.

Step 3: Update the data logging process to record suggestion probabilities together with customer choices.

Step 4: Employ the new system at one of the fitness clubs working with Exerp.

This simple recommendation system will build the basis for data collection and embedding of more sophisticated personalization strategies that were discussed earlier.

Research methods

There are two basic research questions underlying the research proposal: (1) design of algorithms that can handle both stochastic and adversarial environments and (2) design of algorithms that can filter side information. The two research questions come back in different incarnations in the two subprojects that we focus on – recommendation of fitness classes and engineering of personal discounts for snacks and drinks. Below we provide a brief description of the key approaches to these two questions.

Handling stochastic and adversarial environments

The heart of online learning is the exploration-exploitation trade-off. It is the trade-off between playing the action that was the best so far (exploiting) and trying an action that was so far suboptimal or unexplored (exploring). It is important to explore, because our estimate of "the best action" may be misleading due to statistical fluctuations. However, it is also important not to over-explore, because during exploration we play actions that are potentially suboptimal and, therefore, exploration costs in performance. In classical approaches, the exploration-exploitation trade-off is treated differently in stochastic and adversarial environments. In the adversarial regime the algorithms are more "careful" and explore more, because the adversary can change the best action at any time. In the stochastic regime the algorithms are more "aggressive" and reduce exploration when significant gap between performance of the best action compared to all other actions is observed. The reduction of exploration allows improving the performance of the algorithm, but such reduction is risky in the adversarial regime. In a recent work we designed the EXP3++ algorithm [SS14], which combines the careful approach of the celebrated EXP3 algorithm for adversarial bandits [ACB+02] with a new exploration strategy that allows to exploit gaps between performance of different actions.

One playing strategy for adversarial partial monitoring games is based on an algorithm that is very similar to the EXP3 algorithm for limited feedback games [CBL06]. The first research task of the proposal is to generalize the exploration strategy of the EXP3++ algorithm to partial monitoring games.

Filtering the relevant side information

Playing strategies with side information (both bandits and partial monitoring games) can be seen as a mapping from the side information bits to the actions. The more side information bits are relevant the more complex the optimal mapping is. On the other extreme, if all bits are irrelevant the optimal mapping is simple and can ignore the side information altogether. Filtering of relevant side information in bandits with side information is based on combination of an EXP3-like strategy with PAC-Bayesian analysis [SAL+11]. PAC-Bayesian analysis [McA98, SLCB+12] is a generalization of a union bound. The algorithm implicitly enumerates all possible mappings from side information to actions and in the generalized union bound gives preference to "simple" mappings (those that use less side information bits for picking the action). This preference (expressed through the prior distribution over all possible mappings) makes identification of simple mappings much faster (and complex mappings take longer time to learn in any case, so it does not degrade the performance of the algorithm in this regard). We note that the algorithm does not enumerate the mappings explicitly (which would take exponential time) and is, therefore, very efficient in terms of run time.

The second research task of the proposal is to generalize this side information filtering strategy to partial monitoring games with side information.

The third research task of the proposal is to combine information filtering technique with exploration strategy of EXP3++, both for limited feedback and for partial monitoring games.

All research parts of the proposal will be carried out at the University of Copenhagen under supervision of Assistant Prof. Yevgeny Seldin and Prof. Christian Igel. Implementation and application of the resulting algorithms will be carried out at Exerp with support from Laurent Czinczenheim.

E. Publication plan

The intention is to publish two papers with theoretical results at the top tier machine learning conferences and two accompanying papers presenting the results from the application perspective at conferences on recommender systems and fitness industry.

The first theoretical paper will present the results in personalized group fitness classes recommendation and the second the results in personalized recommendation of over-the-counter products. The first theoretical paper will be submitted for publication toward the end of the second year of PhD and the second paper will be submitted for publication toward the end of the third year of PhD. The papers will be submitted to Advances in Neural Information Processing (NIPS), International Conference on Machine Learning (ICML), and/or Artificial Intelligence and Statistics (AISTATS) conferences, depending on the match between the project's and the conference's deadlines.

The accompanying application papers will be submitted to the ACM Conference Series on Recommender Systems (RecSys), The International World Wide Web Conference (WWW), ACM Transactions on Intelligent Systems and Technology, and/or to the Journal of Fitness Research.

F. Courses, conferences and stays abroad

Courses

- Business Course for Industrial PhD students, 7.5 ECTS, DTU, Fall 2015
- Responsible Conduct of Research (mandatory course for PhD students at University of Copenhagen), 1 ECTS, University of Copenhagen, Fall 2015
- Academic Writing, 2 ECTS, The PhD School of Science, University of Copenhagen, Spring 2016
- Online and Reinforcement Learning Summer School, 2.5 ECTS, University of Copenhagen, 29.06.2015 – 3.07.2015
- Advanced Topics in Machine Learning (course), 7.5 ECTS, University of Copenhagen, Block 1, 31.08.2015 – 08.11.2016
- DIKU Image Group PhD Summer School 2015, 2.5 ECTS, University of Copenhagen
- DIKU Image Group PhD Summer School 2016, 2.5 ECTS, University of Copenhagen
- DIKU Image Group PhD Summer School 2017, 2.5 ECTS, University of Copenhagen
- MLSS – Machine Learning Summer School 2016, 2.5 ECTS, see http://www.mlss.cc

Conferences

The student will attend at least two machine learning conferences. The conferences will be selected from the top tier conferences in the field: Neural Information Processing Systems (NIPS), International Conference on Machine Learning (ICML), Artificial Intelligence and Statistics (AISTATS), and Conference on Learning Theory (COLT). The student will also attend at least one conference on recommender systems from either the ACM Conference on Recommender Systems (RecSys) or the International World Wide Web Conference (WWW). Furthermore, the student will attend at least one fitness industry conference/seminar/fair to present the results to potential Excerp business partners at the end of the project.

Stays abroad

The candidate will have a possibility of spending 3 months abroad at other universities or research institutions. The tentative plan is to spend 2 months in one lab during the second year for in-depth collaboration and to make a 1-month tour through 2-3 labs during the third year to present the results and exchange ideas. Preliminary hosts for the stays abroad include Prof. Nicolò Cesa-Bianchi at Università degli Studi di Milano (Milan, Italy) and Prof. Claudio Gentile at Universita' dell'Insubria (Varese, Italy); Prof. Shie Mannor and Prof. Koby Crammer at The Technion (Haifa, Israel); Prof. John Shawe-Taylor at University College London (UK); and Sebastien Bubeck, Ofer Dekel, Lihong Li, Alex Slivkins, and John Langford at Microsoft Research (Redmond and New York City, USA).

G. Dissemination plan

During the project, the scientific results will be disseminated through the following activities:

- Writing and publication of papers and posters: 140 hours
- Participation in conferences and workshops with presentations: 60 hours
- Presentations of research at Excerp, DIKU, and to Excerp’s business partners: 30 hours
- Contributions to machine learning related courses at DIKU (Machine Learning, Advance Topics in Machine Learning): 40 hours
- Business course and report: 120 hours
- Making source code available to the community: 70 hours

General-purpose algorithms will be made freely available as open source. Thereby the project will serve the scientific community at large. For example, the algorithms could be added to the open source machine learning software library Shark [IHMG08], which is maintained at DIKU's MLLab.

H. Project schedule

Below we provide the project schedule by semesters. For each semester we indicate the percentage of time the candidate will spend at the company and at the university (stays abroad are counted for the university time). Please, refer to Section D for detailed description of the work packages.

Semester 1 (University – 70%; Company – 30%):

- Studying the basics of online learning
- Attending "Online Learning Summer School" and "DIKU Image Group PhD Summer School"
- Attending "Advanced Topics in Machine Learning" course
- Attending the "Business Course for Industrial PhDs" and the "Responsible Conduct of Research" course
- Getting access to Exerp data and setting up a basic work environment (starting WP0)

Semester 2 (University – 30%; Company – 70%):

- Performing literature review on topics directly related to the project
- Attending course "Academic Writing"
- **Milestone 1:** finishing WP0
- Starting WP1

Semester 3 (University – 50%; Company – 50%):

- Continuing the work on WP1
- Attending "DIKU Image Group PhD Summer School"
- First stay abroad (2 months)

Semester 4 (University – 40%; Company – 60%):

- Finishing WP1
- **Milestone 2:** writing the first theoretical and the first application papers based on the results of WP1
- Starting WP2

Semester 5 (University – 50%; Company – 50%):

- Finishing WP2
- **Milestone 3:** Writing the second theoretical paper based on the results of WP2
- Attending "DIKU Image Group PhD Summer School"
- Traveling to two conferences to present the results of WP1

Semester 6 (University – 60%; Company – 40%):

- Tour abroad (1 month, visiting 2-3 labs and presenting the results)
- Writing application oriented paper on using recommender systems and online learning in the fitness industry
- Traveling to one or two conferences to present the results of WP2
- Writing the thesis

I. Time allocation

Allocation of the Industrial PhD candidate's time	in months	in % of project time
In Danish division of host company	18	50%
In non-Danish divisions of host company	-	-%
At other companies or organisations	-	-%
At the host university	15	42%
At other research institutions	3	8%

J. Company

The company and its activities

Exerp is a Copenhagen based software development company. Exerp is today employing 35 people, and is solely owned by current and old staff of the company. Exerp was founded in 1994 by two biology students, and has kept growing ever since. The last three years the company has increased the revenue with min 25% each year, expecting a revenue in 2014 of DKK 40 million. The primary income of Exerp is SAAS, which the company provides for its customers in the fitness and leisure industry around Europe. Besides the SAAS, Exerp earnings come from consultancy and hardware.

Most of the Exerp employees have university degrees. However, Exerp as a company has only done little research work before this project. For instance, Exerp has conducted analyses of user behaviour amongst customer staff, to evaluate modifications in user interfaces and assisted customers in integration of different data mining and retention tools.

The candidate's placement in the company

The PhD candidate will work directly under Laurent Czinczenheim, the COO of Exerp. In the department, the PhD will sit together with the rest of the business intelligence team, who works with data analyses, optimisation and usage. Exerp as a company has a flat management without unnecessary hierarchies, so if needed the different departments often interact and consult each other to solve tasks as swift as possible. The PhD will be included in the same internal processes.

Exit strategy

Exerp is a stable company with a broad base of customers and is confident in its ability to support the project until the end of PhD candidate's education.

K. University

University of Copenhagen

With over 38,000 students and more than 9,000 employees, the University of Copenhagen (UCPH) is one of the largest institutions of research and education in the Nordic countries. The university was founded in 1479. Today it comprises six faculties offering over 200 study programs.

The University of Copenhagen is constantly ranked among the best universities in Europe. For example, it ranks 39th in the world according to the Shanghai ranking of 2014, also known as Academic Ranking of World Universities (ARWU), making UCPH third on the list of continental European universities. Eight UCPH alumni became Nobel laureates. The University of Copenhagen is a member of the International Alliance of Research Universities (IARU), which consists of ten universities (ANU, ETHZ; NUS; PKU; UC Berkeley; University of Tokyo; Cambridge, Oxford, and Yale University).

Department of Computer Science (DIKU)

The Department of Computer Science at the University of Copenhagen (DIKU) was founded in 1970 with Turing Award recipient Peter Naur being its first professor. It is among the top 50 computer science departments worldwide according to the latest ARWU ranking, currently employing three ACM fellows.

Machine Learning Lab

The PhD will be part of DIKU' Machine Learning Lab (MLLab), which offers a stimulating scientific environment for machine learning research (see http://image.diku.dk/MLLab). The MLLab is part of the *The Image Section*, which is world-famous for its research in the fields of Image Processing, Computer Vision, Computer Graphics, Machine Learning, and Information Retrieval. Thus, the PhD candidate has the opportunity to communicate and collaborate with several PhD students and scientific staff who are working on related problems. The two academic supervisors are machine learning specialists with a track record in online learning (e.g, [GI08, HMI09, SAL+11, SBCA14, SS14, ST10]), in particular Yevgeny Seldin, who is a renowned online learning expert.

The group will provide the computational infrastructure for conducting large-scale experiments, in particular through access to a new 256 core compute cluster. The award-winning open source machine learning library Shark [IHMG08] is maintained at the MLLab and can be utilized within the project.

L. Third parties

Does not apply.

Bibliography

[ACB+02] Peter Auer, Nicoló Cesa-Bianchi, Yoav Freund, and Robert E. Schapire. The nonstochastic multiarmed bandit problem. SIAM Journal on Computing, 32 (1), 2002.

[BCB12] Sébastien Bubeck and Nicolò Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. Foundations and Trends in Machine Learning, 2012.

[BS12] Gábor Bartók and Csaba Szepesvári. Partial monitoring with side information. Algorithmic Learning Theory (ALT), 2012.

[BZS12] Gábor Bartók, Navid Zolghadr and Csaba Szepesvári. An adaptive algorithm for finite stochastic partial monitoring. International Conference on Machine Learning (ICML), 2012.

[CBGZ13] Nicolò Cesa-Bianchi, Claudio Gentile, and Giovanni Zapella. A gang of bandits. In Advances on Neural Information Processing Systems (NIPS), 2013.

[CBL06] Nicolò Cesa-Bianchi and Gábor Lugosi. Prediction, learning, and games. Cambridge University Press, 2006.

[GI08] Tobias Glasmachers and Christian Igel. Second Order SMO Improves SVM Online and Active Learning. Neural Computation, 20, 2008

[CL12] Olivier Chapelle and Lihong Li. An Empirical Evaluation of Thompson Sampling. In Advances in Neural Information Processing Systems (NIPS), 2012.

[GLZ14] Claudio Gentile, Shuai Li, and Giovanni Zappella. Online clustering of bandits. In JMLR Workshop and Conference Proceedings, 32 (ICML), 2014.

[HMI09] Verena Heidrich-Meisner and Christian Igel. Hoeffding and Bernstein Races for Selecting Policies in Evolutionary Direct Policy Search. Proceedings of the International Conference on Machine Learning (ICML), 2009

[IHMG08] Christian Igel, Verena Heidrich-Meisner, and Tobias Glasmachers. Shark. Journal of Machine Learning Research, 9, 2008.

[LCLS10] Lihong Li, Wei Chu, John Langford, and Robert E. Schapire. A Contextual-Bandit Approach to Personalized News Article Recommendation. In Proceedings of the International Conference on World Wide Web (WWW-10), 2010.

[McA98] David McAllester. Some PAC-Bayesian theorems. In the proceedings of the Conference on Learning Theory (COLT), 1998.

[Netflix] Netflix Challenge. http://www.netflixprize.com/

[SAL+11] Yevgeny Seldin, Peter Auer, François Laviolette, John Shawe-Taylor, and Ronald Ortner. PAC-Bayesian analysis of contextual bandits. In Advances in Neural Information Processing Systems (NIPS), 2011.

[SBCA14] Yevgeny Seldin, Peter L. Bartlett, Koby Crammer, and Yasin Abbasi-Yadkori. Prediction with limited advice and multiarmed bandits with paid observations. In JMLR Workshop and Conference Proceedings, 32 (ICML), 2014.

[SLCB+12] Yevgeny Seldin, François Laviolette, Nicolò Cesa-Bianchi, John Shawe-Taylor, and Peter Auer. PAC-Bayesian inequalities for martingales. *IEEE Transactions on Information Theory*, 58(12), 2012.

[SS14] Yevgeny Seldin and Aleksandrs Slivkins. One practical algorithm for both stochastic and adversarial bandits. In JMLR Workshop and Conference Proceedings, 32 (ICML), 2014.

[ST10] Yevgeny Seldin and Naftali Tishby. PAC-Bayesian analysis of co-clustering and beyond. Journal of Machine Learning Research, 2010.