**Course Title: Data Science Projects – Business Practicum**

**Course Number: \_\_\_\_\_\_\_\_\_\_\_**

**Number of credit points\_\_5\_\_**

Mini-Semester: of the Academic Year: 2019

Time: (Day & Hour) \_\_\_\_\_\_\_\_\_\_\_\_\_

**Course Instructor: \_Ido Biger\_(chief data officer)**

E –mail: **\_\_\_\_idobi@elal.co.il\_\_\_**

Telephone: **\_\_\_+972-50-669-9604\_\_\_\_\_**

Meeting time for students: \_\_\_\_\_ by appointment \_\_\_\_\_\_ (or by appointment)

**Teaching Assistant: Amit Rappel\_(lead data scientist)**

E –mail: **amitrappel@gmail.com**

Telephone: **+972-52-677-9595**

Meeting times for students:\_\_\_\_\_\_ by appointment \_\_\_\_\_ (or by appointment)

**Syllabus:** (choose and get into details of the relevant parts)

**Course Objectives:**

As part of the on growing need for the business & the academy to join forces, this course was designed.

The Course will allow its participants to execute an end to end data science project for one of the leading companies in Israel.

Students that would successfully finish this course, would experience a Data Science Internship like no other student will.

**Course Content & Scope:**

The students will face a real business challenge, accompanied by the chief data officers of the companies, learning to overcome it by using the relevant data sources, machine learning capabilities and a comprehensive understanding of the business phenomena  
**Teaching Methods:**Weekly Frontal session (on Campus) – Mandatory attendance.

One kick off meeting – client’s site

On site work (businesses HQ) & Meetings – TBD

One wrap up meeting – client’s site

**Teaching Materials:**

Course lectures and walk through guidance.

The students will be divided to teams – each team would be responsible for delivering the E2E project for its customer.

**Readings (Compulsory / Recommended/pre-Requisite):**

* Machine learning fundamental
  + [An Introduction to Statistical Learning](http://www-bcf.usc.edu/~gareth/ISL/) (free book, recommended)
* Python / R – intermediate level
  + [Python for Data Analysis](https://www.oreilly.com/library/view/python-for-data/9781491957653/) (recommended)
  + [Hands-On Machine Learning with Scikit-Learn and TensorFlow](https://www.amazon.com/Hands-Machine-Learning-Scikit-Learn-TensorFlow/dp/1491962291) (recommended)
* Data engineering knowledge (BI)
* Big data fundamentals.

**Student Assessment:**Please indicate the relevant components:

Course Assignments

Throughout the course – **students will present their findings and progress on a weekly basis, working closely with the Course Assistant and the customer’s local team.** 15 minutes presentation by each team every week (from the 3rd session on).

The team would be held responsible for a tight schedule progress alongside a very real-time experience of an actual data team.

Important information:  
How work is to be submitted (hard copy / **E-mail attachment**).  
Length of work: the final wrap up work would be presented as a power point to the customer and with an approximately 10-20 pages of detailed design.   
Submission date. 3 weeks after EOS  
 **Course Plan**

**Lesson 1. Date**

Business lesson 1:

Lecture – 60 minutes by Ido.

60 minutes lecture by Amit.

ML problem types & Measuring model performance

Students Introduction + team workshop.

Teams selection and business matching.

By the end of the first day, the students would be able to learn about the relevant company and prepare themselves to the opening meeting with the business customer.

Each group will be set to prepare the opening presentation for the on-site meeting with the customer.

The presentations will be prepared as part of the workshop.

Each team would hold responsible to learn and prepare itself about the business world of it selected customer.

**Lesson 2. On site meeting (customer site)**

Business lesson 2:

**Kick Off Meeting**

The meeting will take place on the customer’s site. With the business sponsor and his local team. Our groups would present the project / course (Ido will take part in those meetings) and the meeting preparations would be demonstrated (and its quality) while having the conversation with the customer. The outcome of this session would be a meeting summary + a well-defined business problem + Key success factors of the project.

**Lesson 3. Date**

End of the business phase:

Business lesson 3:

The session would deal with the readiness of the teams towards the business challenge they've faced at the customers site. At the end of the session the team would have understand the needs / requirements needed to face the challenge.

By the end of this session – the teams would come up with the full scope of the project as far as needs, customers, relevant sources etc.

90 minutes lecture by Amit.

Feature engineering (prerequisite – pandas)

Business summary of the challenges – teams review and further guidance of the technical issues – by Amit.

This includes the following ML aspects:

* Framing of the business problem
* Defining the relevant evaluation measures
* Mapping the available data and its relevancy

**Lesson 4. Date**

Data Exploration + Data Preparation + generating the data set + understanding the data sources and architecture tools

30 minutes lecture by Ido

60 minutes lecture by Amit

Data resources and pipelines

Facing the technical challenges – data preparation phase.

Learning how the take over the data gaps and data needs to confront with the business issue.

**Lesson 5. Date**

Technical lesson 2:

60 minutes lecture by Amit.

Model selection (prerequisite: scikit-learn)

Data preparation + Model choosing

Technical walkthrough with Amit.

Understanding the data gaps and challenges.

By the end of this session the teams should have:

* A ready-to-use version of their data
* An extensive collection of engineered features
* Good familiarity with the technologies used by the customer

**Lesson 6. Date**

End of phase 2:

Technical lesson 3:

60 minutes lecture by Amit.

Advanced model selection – ensemble methods and hyper-parameters optimization

Modelling and Execution – training and testing several models, choosing the relevant.

Data preparation + Model choosing + solution development – model training

Running and executing the data project – understanding the main gaps and outcomes. Working on summarizing the results and wrapping up the suggestions of the technical phase.

**Lesson 7. Date**

60 minutes lecture by Amit.

60 minutes lecture by Ido

Business summary – preparation for the wrap up meeting – project summary – suggested next steps (business and technical). project review and business summary.

By the end of this session the team should have a working version of their project pipeline (in Scikit-Learn. Not in client’s technology)

**Lesson 8. On site meeting (customer site)**

Wrap up meeting with the customer.

Project conclusions.

This is where it all comes together. The customer sees the final deliverables, the suggested next steps (internally), the project process, HLD, and business issues that the team confronted.