Course program and reading list

Semester 2 Year 2021

School: Adelson School of Entrepreneurship B.A

Data Science Implementation for Entrepreneurs

Lecturer:
Dr. David Movshovitz  dmovshovitz@idc.ac.il

Tutors:
Mr. Yaniv Levi  ylevi@post.idc.ac.il

Teaching Assistant:
Mr. Yaniv Levi  ylevi@post.idc.ac.il

Course No.: 26006  Course Type: Lecture  Weekly Hours: 3  Credit: 3

Course Requirements: Final Paper  Group Code: 212260063  Language: English

Prerequisites

Students who took one of the courses listed below will not be allowed to register to the course Data Science Implementation for Entrepreneurs (26006):

3078 – Data Science- group 1
3411 – Data Science Implementation for Entrepreneurs
“Data are becoming the new raw material of business” (Craig Mundie, Senior advisor to Microsoft CEO). Many business problems in the financial, marketing, advertisement, management, healthcare, etc. require data analytics tools to analyze the massive amount of data collected by the organization, and transforming the raw data into meaningful business insights. In the course we will discuss methods to analyze data and make recommendations such as Naïve Bayes, as well as clustering technologies like K-mean, and will learn how to use them to translate the data into recommendations for solving business problems.

**Lecture topics (tentative):**

<table>
<thead>
<tr>
<th>Lecture #</th>
<th>Lecture Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Using data science for business management – part 1</td>
</tr>
<tr>
<td>2</td>
<td>Using data science for business management – part 2</td>
</tr>
<tr>
<td>3</td>
<td>Introduction to predictive analytics and classifiers</td>
</tr>
<tr>
<td>4</td>
<td>Bayes Theorem</td>
</tr>
<tr>
<td>5</td>
<td>Naïve Bayes Classifiers</td>
</tr>
<tr>
<td>6</td>
<td>Using Naïve Bayes in the real-world</td>
</tr>
<tr>
<td>7</td>
<td>Introduction to clustering</td>
</tr>
<tr>
<td>8</td>
<td>K-Means clustering</td>
</tr>
<tr>
<td>9</td>
<td>Hierarchical Agglomerative Clustering (HAC)</td>
</tr>
<tr>
<td>10</td>
<td>Clustering Summary</td>
</tr>
<tr>
<td>11</td>
<td>Introduction to Recommendation engines</td>
</tr>
<tr>
<td>12</td>
<td>Guest lecture</td>
</tr>
<tr>
<td>13</td>
<td>Other Recommendation techniques</td>
</tr>
</tbody>
</table>
The course will provide the students understanding of:

- The importance of data science to management in the digital era
- The sources of data that can be used for management decisions
- How data is translated into information and meaningful business insights
- The difference between supervised learning and unsupervised learning
- The concept of predicative analytics
- The concepts of classifiers
- The Bayes theorem
- The Naïve Bayes algorithm and its usage for classifiers
- The concepts of clustering algorithms and their usage
  - the parameters that have to be defined for a clustering algorithm
  - their effect on the clustering algorithm results
- The K-mean clustering algorithm, its advantages and weaknesses
- The Hierarchical Agglomerative Clustering (HAC) algorithm, its advantages and weaknesses
- The concepts and main techniques of recommendation systems

---

**Grading**

The course final grade will be composed from:

- final project (at the end of the semester) - 50%
- home assignments during the semester - 25%
- quizzes (during the semester) - 25%

---

**Learning Outcomes**

At the end of the course the student should be able to:

- Explain the importance of data science to management in the digital era
- Identify the sources of data that can be used for management decisions
- Understand how data can be translated into meaningful business insights
- Explain the difference between supervised learning and unsupervised learning, and when we should use each type of machine learning
- Explain concepts of predicative analytics and classifiers
- Explain the Bayes theorem
- Use Naïve Bayes for classification and prediction
- Explain the main techniques used for clustering
- Use the K-mean algorithm to cluster entities
- Explain the main differences between K-Mean clustering and HAC (including the advantages and disadvantages for each these algorithms)
- Use clustering to solve business problems
- Explain the main techniques used for recommendation engines
Lecturer Office Hours

Sunday after class (from 11:30 to 12:30)

Use can contact me by email: dmovshovitz@gmail.com

Additional Notes

Classes will consist of lectures and discussions.

Students are required to attend 80% of classes and any assigned supplementary activities related to the course.

Students are required to submit all the home assignments

Students are required to attend all the quizzes (during the semester)

Reading List

- Galit Shmueli, Nitin R. Patel, and Peter C. Bruce, Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner (English), Wiley
- Provost, Foster, and Tom Fawcett. Data Science for Business: What you need to know about data mining and data-analytic thinking. O'Reilly Media, Inc. 2013, Chapter 9

Additional articles will be uploaded to the course website in moodle.