

מודלים לניבוי 218.5220
החוג לניהול מידע וידע
אוקטובר-דצמבר 2019
מרצה: דייוויד בודוף dbodoff@univ.haifa.ac.il
עוזרת הוראה: דינה בבלי din.bavli@gmail.com

מטרת הקורס היא להפוך את הידע התיאורטי שצבר הסטודנט בשיטות וכלים בתחום מדעי הנתונים, ליכולת מעשית יותר. נתמקד ביכולת לבנות מודל ניבוי בשימוש שיטות של למידת מכונה, שהוא אחד השימושים העיקריים של מדעי הנתונים. הקורס ידגיש התנסות מעשית משולבת בשיעורים פרונטליים. הסטודנט יעבור את כל שלבי הניתוח כדי לפתור בעיה עסקית על ידי בניית מודל ניבוי.

מטלות

כל שבוע תהיה מטלה להגשה. הסטודנטים מוזמנים לעזור אחד לשני, אבל מטלות מוגשות על ידי כל סטודנט בנפרד. בסוף הקורס יינתן מטלה אחת גדולה יותר. אין בחינה

מרכיבי ציון

5 שיעורי בית, 10% כל אחד
מטלה מסכמת, 50%

נושאים

Business Understanding

Data Understanding (Exploratory Data Analysis)

Some important concepts: base table, distributions, outliers, missing values

Data Preparation

Some important concepts: normalization, binning, handling outliers, handling missing values

Feature Selection and Feature Engineering

Some important concepts: PCA/SVD, information gain, correlation

Modeling

Decision Trees (for discrete target. Attention paid to overfitting)

Regression Trees (for continuous target)

k-nearest-neighbor (and back to feature selection, missing values, normalization)

naïve Bayes (with smoothing, binning for continuous features)

Linear Regression (gradient descent)

Evaluation

בסוף הקורס הסטודנט יידע לבחון את ההיתכנות, לתכנן, ולממש פתרון לבעיה עסקית, על ידי בניית מודל ניבוי בשימוש שיטות של למידת מכונה.