למידה חישובית – תרגיל 3

למציאת הפיצ'רים המשפיעים ביותר, השתמשנו ב Mutual Information. בהמשך לתרגיל הקודם, השוונו בין 2 שיטות למציאת מינימום:

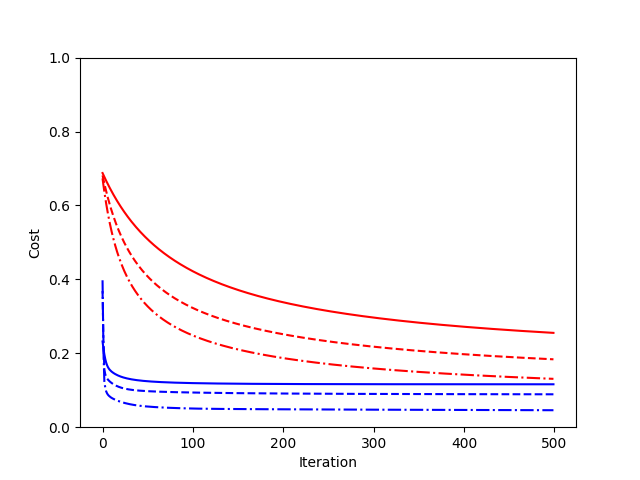
1. BGD (Batch Gradient Decent)
2. SGD (Stochastic Gradient Decent)

בשתי הניסויים השתמשנו בפרמטרים הבאים:

1. Alpha = 0.01
2. Max number of iterations = 500
3. Acceptable cost threshold = 0.01
4. Training set = 500

השוונו בין 3 הרצות, כך שבכל אחת השתמשנו במספר שונה של הפיצ'רים הכי משפיעים.

גרף של ירידת השגיאה ככל שהתקדמנו:



באדום – עקומת ירידת השגיאה של BGD

בכחול – עקומת ירידת השגיאה של SGD

קו רצוף (\_\_\_): שימוש ב5 הפיצ'רים המשפיעים ביותר

קו מקווקו (---):שימוש ב15 הפיצ'רים המשפיעים ביותר

קו נקודה (-.-.-.):שימוש ב30 הפיצ'רים המשפיעים ביותר (כל הפיצ'רים)

**תוצאות הניסוי:**

|  |  |  |  |
| --- | --- | --- | --- |
| 30 Features | 15 Features | 5 Features | TRAIN ERROR |
| 2.8% | **5.8%** | **7%** | **BGD** |
| 1.2% | **3.4%** | **4.8%** | **SGD** |

|  |  |  |  |
| --- | --- | --- | --- |
| 30 Features | 15 Features | 5 Features | TEST ERROR |
| 2.8% | **0%** | **1.45%** | **BGD** |
| 1.45% | **1.45%** | **1.45%** | **SGD** |