

האוניברסיטה העברית בירושלים

סילבוס

ביג דאטה ובינה עסקית למדע המדינה - 56146

תאריך עדכון אחרון 08-03-2023

נקודות זכות באוניברסיטה העברית: 4

תואר: בוגר

היחידה האקדמית שאחראית על הקורס: מדע המדינה

השנה הראשונה בתואר בה ניתן ללמוד את הקורס: 0

סמסטר: סמסטר ב'

שפת ההוראה: עברית

קמפוס: הר הצופים

מורה אחראי על הקורס (רכז): אלון פלד

דוא"ל של המורה האחראי על הקורס: peled.alon@gmail.com

שעות קבלה של רכז הקורס: שני 1300-1430

מורי הקורס:

תאור כללי של הקורס:

האם וכיצד אפשר לנתח סרטי וידאו ותמונות כדי לחזות אם ראש מדינה משקר לאומה? כיצד ניתן לנתח רמזים טקסטואליים בנאומים של חברי וחברות כנסת בשנת כהונתם הראשונה כדי לנבוא אם הם יהפכו למחוקקים מצטיינים? מהפכת נתוני העתק ("ביג דאטה") והבינה המלאכותית מאפשרת לענות על שאלות שכאלו דרך הזרמת מידע קולי, חזותי וטקסטואלי מגוון ל"בריכות מידע" עצומות, היתוך המידע בבריכות אלו, הפעלת מודלים של למידת מכונה ולמידה עמוקה, והפקת תובנות חדשות. הופעתם של עננים מסחריים דוגמת אלו של גוגל, מיקרוסופט, אמזון ויבמ מעניקה לנו אפשרות זולה, מהירה ופשוטה לפתח תשתיות וממשקי משתמש כדי לענות לשאלות אותן לא העזנו לשאול אתמול

מטרות הקורס:

מטרת הקורס היא להקנות לתלמידים הבנה של מושגי יסוד בתחום הביג דאטה ובינה מלאכותית. לבחון את הרקע, הסיבות, וההיסטוריה דרכה מושגים אלו התפתחו. להבין את האתגרים, הקשיים, ויחסי התן-קח ("טרייד-אוף") הכרוכים בעבודה עם כלי תוכנה המיישמים מושגים אלו. ובעיקר - לחוות התנסות ראשונית מלהיבה של עבודה יצירתית עם כמויות עצומות של נתונים בענן מסחרי כדי לשכנע את התלמידות והתלמידים לנסות ולהמשיך לעבוד עם ביג דאטה ובינה מלאכותית בעצמם.

תוצרי למידה

בסיומו של קורס זה, סטודנטים יהיו מסוגלים:

בסיומו של הקורס, הסטודנטים יהיו מסוגלים:

להגדיר מושגי יסוד כמו ביג דאטה, למידת מכונה, והמחשה ויזואלית של נתונים.

לכרות נתונים על ידי כתיבת קוד ושימוש בענן מסחרי אחד.

לייצר ולהפעיל מודלים מבוססי ML-Auto לניתוח טקסט, תמונה ווידאו.

למצוא תשובות לבעיות תכנות שונות בכוחות עצמם על ידי הפעלת כלים מסוג chatgpt

stackoverflow.

ללמוד בכוחות עצמם מושגי ביג דאטה ובינה מלאכותית שונים תוך שימוש במקורות מקוונים.

להכיר (היכרות ראשונית) מושגים מתקדמים דוגמת בינה למידה עמוקה.

דרישות נוכחות (%) :

90%

שיטת ההוראה בקורס: שני 13:00-14:45 - 26502

חמישי 08:30-10:15 - 26502

רשימת נושאים / תכנית הלימודים בקורס:

פגישה 1 (שני 13 מרץ):

הקורס: יעדים ותכנית ההדרכה. היכרות. חלוקה לקבוצות (לא יותר מארבעה תלמידים בצוות!).
קריאה: אין

פגישה 2 (חמישי 16 מרץ):

ברוכים הבאים לעולם הנפלא של ביג דאטה ובינה מלאכותית!
(בחירת פרויקטים צוותיים)

קריאה: 2013 ,Data Big ,Schönberger-Mayer
עיון:

<https://blog.gdeltproject.org/>

פגישה 3 (שני 20 מרץ)

“עשרת הדיברות” לצעדים ראשונים בעולם הנתונים
קריאה: 2016 Destruction Math of Weapons .Neil'O
עיון:

<https://teachablemachine.withgoogle.com/>

פגישה 4 (חמישי 23 מרץ):

היכרות עם כלים מרכזיים בקורס והתקנות
Notepad++, chatgpt, Stackoverflow, Google Cloud Platform
קישורים

<https://openai.com/blog/chatgpt>
<https://notepad-plus-plus.org/downloads/>
<https://stackoverflow.com/>

פגישה 5 (שני 27 מרץ):

מושגי יסוד בעולם פלטפורמת הענן (platform cloud)

פגישה 6 (חמישי 30 מרץ):

פגישה ראשונה עם מאגר ביג דאטה - ג'דלט (Gdelt)
קריאה: 2013 GDELT ,Schrodte & Leetaru

חופשת פסח - 2 - 15 אפריל

פגישה 7 (שני 17 אפריל):

פתיחת חשבונות ענן אישיים בפלטפורמת הענן של גוגל והיכרות ראשונית עם כלי הענן של גוגל

פגישה 8 (חמישי 20 אפריל):

Google bigquery "Hello World" SELECT...FROM... WHERE...ORDER BY ...LIMIT
קריאה: 2004 Basics Database Relational ,Sato & King

חופשת יום הזיכרון ויום העצמאות - 24 - 26 אפריל

פגישה 9 (חמישי 27 אפריל):

Google bigquery GROUP BY ...ORDER BY

פגישה 10 (שני 1 מאי):
Google bigquery Working with Arrays (UNNEST... ARRAY_AGG)

פגישה 11 (חמישי 4 מאי):
Google bigquery JOINS 1

פגישה 12 (שני 8 מאי):
Google bigquery JOINS 2

פגישה 13 (חמישי 11 מאי):
Google bigquery - Types and Type-Functions

פגישה 14 (שני 15 מאי):
Google bigquery - Advanced Functions

פגישה 15 (חמישי 18 מאי):
Google bigquery - Analytical Functions

פגישה 16 (שני 22 מאי):
Google bigquery - Creating a Data Pipeline

חופשת שבועות - 25 - 26 מאי

פגישה 17 (שני 29 מאי):
Google bigquery - Scheduling the Data Pipeline With Partitioning

פגישה 18 (חמישי 1 יוני):
Google Data Looker - Creating a Simple Dashboard
קריאה: 2001 Excellence Graphical, Tufte

פגישה 19 (שני 5 יוני):
Google Data Looker - Displaying Timeline, Trend, and Geography

פגישה 20 (חמישי 8 יוני):
Google Cloud Videointelligence API - Analyzing photo
קריאה: 2018 Background, Algorithms Learning Machine Master, Brownlee
עיון:

<https://machinelearningmastery.com/>
<https://www.kaggle.com/learn/intro-to-machine-learning>
<https://carto.com/help/tutorials/getting-started-with-carto-builder/>

פגישה 21 (שני 12 יוני):
Google Cloud Videointelligence API - Analyzing video

פגישה 22 (חמישי 15 יוני):
Google Cloud Auto-ML - Creating a Model to Analyze Photos

פגישה 23 (שני 19 יוני): עבודה מורה-צוותים על הפרוייקטים

פגישה 24 (חמישי 22 יוני): עבודה מורה-צוותים על הפרוייקטים

פגישה 25 (שני 26 יוני): מצגות פרוייקטים א'

פגישה 26 (חמישי 29 יוני): מצגות פרוייקטים ב'

חומר חובה לקריאה:

Booch, Grady. *Object-oriented analysis and design with applications*. Redwood City, CA: The Benjamin/Cummings Publishing Company, Inc. 2007.
Brownlee, Jason. *Master Machine Learning Algorithms - Discover How They Work and Implement Them From Scratch*. 2018. Available at:
<https://machinelearningmastery.com/master-machine-learning-algorithms>.

King, Gary., Keohane, Robert. O., & Verba, Sidney. *Designing social inquiry: Scientific inference in qualitative research*, Princeton: Princeton University Press, 1994.

King, Ro and Sato, Todd. "Relational Database Basics" *Information Management* -(2004). Available at <https://www.information-management.com/news/relational-database-basics>.

Leetaru, Kalev and Schrod, Philip A. "GDELT: Global Data on Events, Location and Tone, 1979-2012. 2013. Available at
<http://data.gdeltproject.org/documentation/ISA.2013.GDELT.pdf>.

Mayer-Schönberger, Viktor, and Kenneth Cukier. *Big Data: A Revolution That Will Transform How We Live, Work and Think*. London: John Murray, 2013.

O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown Publishers, 2016. 272p.

Tufte, Edward R., 1942-. *The Visual Display of Quantitative Information*. Cheshire, Connecticut: Graphics Press, 2001.

חומר לקריאה נוספת:

Brownlee, Jason. *Master Machine Learning Algorithms - Discover How They Work and Implement Them From Scratch*. 2018. Available at:
<https://machinelearningmastery.com/master-machine-learning-algorithms>.

King, Ro and Sato, Todd. "Relational Database Basics" Information Management -(2004). Available at <https://www.information-management.com/news/relational-database-basics>.

Leetaru, Kalev and Schrod, Philip A. "GDELT: Global Data on Events, Location and Tone, 1979-2012. 2013. Available at <http://data.gdeltproject.org/documentation/ISA.2013.GDELT.pdf>.

Mayer-Schönberger, Viktor, and Kenneth Cukier. Big Data: A Revolution That Will Transform How We Live, Work and Think. London: John Murray, 2013.

O'Neil, Cathy. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown Publishers, 2016. 272p.

Tufte, Edward R., 1942-. The Visual Display of Quantitative Information. Cheshire, Connecticut: Graphics Press, 2001.

הערכת הקורס - הרכב הציון הסופי :

מבחן מסכם בכתב/בחינה בעל פה 0 %
הרצאה 0 %
השתתפות 10 %
הגשת עבודה 0 %
הגשת תרגילים 30 %
הגשת דו"חות 0 %
פרויקט מחקר 60 %
בחנים 0 %
אחר 0 %

מידע נוסף / הערות:

הערות חשובות:

תלמיד שייעדר מיותר משלוש פגישות מעבדה ללא הסבר מספק לא יקבל ציון עבור הקורס כל תלמיד/ה יתבקש להגיע לפגישה אישית עם המרצה במהלך הסמסטר.

כל תלמיד/ה נדרש/ת להגיש שני תרגילים. כל תרגיל יוגש למורה שבוע לאחר שהתלמיד/ה מקבל/ת התרגיל.

התלמידים יתבקשו להתקין תוכנות חינוכיות שונות על המחשבים האישיים שלהם. את ההתקנות נבצע ביחד במעבדות הראשונות בקורס.

התלמידים ידרשו לפתוח חשבון אישי בשירותי ענן (למשל: Platform Cloud Google-GCP) תוך שימוש בפרטים אישיים, חשבון ג'מייל לא אוניברסיטאי, וכרטיס אשראי אישי. השימוש במקורות GCP חנים אך ורק אם התלמידים עוקבים בקפדנות אחר הוראות המורה. כל שימוש במקורות GCP שלא על פי הנחיות המורה עשוי לעלות כסף ויהיה על חשבון התלמיד! את חשבון פלטפורמת הענן של גוגל

אתם חייבים לפתוח באמצעות חשבון ג'מייל פרטי שלכם (לא תוכלו להשתמש בחשבון הג'מייל של האוניברסיטה העברית למטרה זו). את החשבונות נפתח ביחד במעבדה הראשונה בקורס (בבקשה לא לפתוח חשבונות אלו לפני כן כיוון שהקרדיט חינם לשימוש בהם מוגבל לשלשה חודשים!). אפשר, כמובן, לפתוח חשבון ג'מייל נפרד אך ורק למטרת השימוש ב GCP אולם גם אז תדרשו למסור לגוגל פרטי כרטיס אשראי אישי כדי לפתוח את החשבון. תלמיד או תלמידה אשר מסיבה כלשהי לא יכול לעמוד בתנאי סף זה לא יוכל/תוכל להשתתף בקורס.