

שם הקורס: עיבוד קול למערכות תבוניות

Speech & Voice Intelligent Processing

מתכונת הלימוד: שיעור + מעבדה

שעות שבועיות: 4

נקודות זכות: 3

דרישות קדם: ידע בשפות תכנות (Java, Python), עבודה עם API's ו- Open source coding, יכולת לימוד עצמי.

מטרות הקורס:

"עיבוד קול למערכות תבוניות" הוא קורס מתקדם אשר מציע הבנה תיאורטית משולבת בהבנה מעשית של איך הדיבור האנושי יכול להיות מעובד על ידי מחשב. התחומים שאנו מכסים הם: זיהוי דיבור ASR, סינתזה דיבור TTS, זיהוי דובר ביומטרי, זיהוי רגש דיבור, ניתוח שפה בדיבור ועוד.

אחד מתחומי הבינה המלאכותית הוא עיבוד קול ודיבור. בשנים האחרונות הוא מהווה גם מרכיב חיוני בכל ממשק אדם-מכונה כחלק מיישומים/אפליקציות, מוצרי IoT, מיחשוב לביש ורובוטים. תחום חדש ומתפתח שנקרא Conversational Interaction מחייב ידע של תוכנות שהן Voice enabled.

תיאור הקורס:

הקורס מיועד לסטודנטים המבקשים להבין כיצד לעבד נתונים אמיתיים ליישומים אמיתיים, תוך יישום טכניקות סטטיסטיות ומכונות למידה, כמו גם עבודה עם אתגרי הטכנולוגיה. עבודה זו תהיה מבוססת על ערכת כלים קיימים SDK של מוצרים בשילוב עם פיתוח ואינטגרציה אישית. סטודנטים צפויים לעבוד באופן עצמאי בצוותים כדי לחקור וליישם יישום עיבוד דיבור או מודול אשר באחד התחומים המתוארים מעלה. הם יידרשו ליישם את השיקולים התיאורטיים לתוך יישום/שימוש אמיתי אשר ידרוש שילוב של כלים קיימים בצורה חדשה ליצור (POC) proof of concept אשר יוגש ויודגם בכיתה. התחומים שנציג הם:

Speech recognition Speech Synthesis Speech Biometrics Emotion detection
Speech Analytics Language Identification Conversational Interaction

נושאים שיילמדו בקורס:

1. מונחים עיקריים והגדרות כמו גם נושאים תיאורטיים כגון Acoustics Phonetics ו-Linguistic levels.

2. בעיות מעשיות של עיבוד ואוטומציה ויישום באמצעות שימוש בכלי עיבוד ועיצוב כגון PRAAT, TOBI, MATLAB

3. יישומים עם ערכות (SDK) לפיתוח מוצרי דיבור של חברות (בשיתוף עם חברות הזנק):
AWS, Google APIs, API.AI, ווטסון,

הציון הסופי יקבע ע"פ הקריטריונים הבאים:

מטלות	דרישה	ציון
תרגילי מעבדה	4 תרגילים במהלך הסמסטר	50 %
פרויקט	הגשת פרויקט קורס ומצגת בע"פ	50 %

ציון "עובר" – 60

נוכחות חובה בשיעורים: 80%

הערה: הקורס כולל אלגוריתמים אוטומטיים ומתודולוגיות ליצירת דיבור, תעתוק ועיבוד. זהו חלק מהכרחי של כל מעצב תוכנה בוגר מדעי המחשב. ניתן לבצע את הקורס כבסיס לקורס NLP לתלמידי שנה ג'. זה יעזור להקל על פרויקטים הדורשים אלגוריתמים ופיתוח תוכנה עם שימוש בשפות התכנות Python, Java, עבודה עם API וקוד פתוח.

ביבליוגרפיה

1. Lawrence R. Rabiner and Ronald W. Schafer (2007), "*Introduction to Digital Speech Processing*", Foundations and Trends® in Signal Processing: Vol. 1: No. 1–2, pp 1-194
2. Benesty, J. Sondhi, M. Huang, Y. Eds. (2008) Handbook of Speech Processing, Springer Publishing.
3. Jurafsky, D. (2000). *Speech & language processing*. Pearson Education India.
4. Pieraccini, R. (2012). The Voice in the machine: Building Computers that understand Speech. MIT Press.
5. List of relevant articles