

סיכום רשתות נוירונים

רשת נוירונים מלאכותית (Artificial Neural Network) או בקיצור ANN הינה דרך להתאים (Fitting) מודל לא לינארי לנתונים. משתני היציאה (Outputs) אינם קשורים במישרין למשתני הכניסה (Inputs). קיימות מספר שכבות נסתרות (Hidden Layers) המעורבות בתהליך שכוללות נוירונים הערכים שבנוירונים בשכבה הנסתרת הראשונה קשורים למשתני הכניסה; הערכים שבנוירונים בשכבה הנסתרת השנייה קשורים לערכים שבנוירונים בשכבה הנסתרת הראשונה; וכך הלאה. משתני היציאה מחושבים מתוך הערכים שבנוירונים בשכבה הנסתרת הסופית (Final Layer).

הפונקציות שמגדירות את הקשרים נקראות פונקציות אקטיבציה (Activation Functions). פונקציית הסיגמואיד (Sigmoid Function), אשר דיברנו עליה בהקשר של רגרסיה לוגיסטית, לעיתים קרובות משמשת כפונקציית אקטיבציה על מנת לקשר בין (א) הערכים שבנוירונים בשכבה הנסתרת בשכבה הנסתרת ה- i לבין (ב) הערכים שבנוירונים בשכבה הנסתרת ה- $i-1$. כאשר אומדים ערכים נומריים או אז פונקציית האקטיבציה שמקשרת בין משתני היציאה לבין הנוירונים בשכבה הנסתרת הסופית היא על פי רוב לינארית. כאשר הנתונים מסווגים, הרי שפונקציית הסיגמואיד מתאימה וראויה יותר עבור השלב האחרון שתואר.

אלגוריתם מורד הגרדיאנט (Gradient Descent) משמש למזעור השגיאה הריבועית הממוצעת (MSE - Mean Squared Error) ברשת נוירונים מסוימת. ניתן לחשוב על חישוב המינימום כעל מציאת התחתית של עמק כלשהו. האלגוריתם עושה צעדים במורד העמק כאשר כל צעד עוקב אחר הקו של המורד התלול ביותר. בחירת הגודל הנכון עבור הצעד, המכונה קצב הלימוד (Learning Rate), הינה היבט חשוב של אלגוריתם מורד הגרדיאנט.

ברגיל, רשתות נוירונים כרוכות בעשרות אלפי פרמטרים. אפילו אם היה אפשר למצוא את ערכי הפרמטרים אשר ממזערים לחלוטין את פונקציית המטרה עבור סט האימון (Training Set), עדיין הדבר לא היה רצוי היות והוא כמעט בוודאות יסתיים בהתאמת יתר (Over-Fitting), בעיה שכחה שקיימת באלגוריתמים של למידת מכונה, שבה האלגוריתם לומד טוב מדי את נתוני סט האימון, אך לא נותן תחזיות נכונות לנתוני סט הבדיקה). בפרקטיקה, כלל עצירה (Stopping Rule) מיושם כך שאימון באמצעות אלגוריתם מורד הגרדיאנט נעצר כאשר התוצאות המתקבלות עבור סט האימון (Validation Set) חורגות מהתוצאות המתקבלות עבור סט האימון.

רשת נוירונים קונבולוציונית (Convolutional Neural Network) או בקיצור CNN הינה רשת נוירונים שבה הנוירונים בשכבה אחת קשורים לתת-קבוצה של נוירונים בשכבה הקודמת במקום לכול הנוירונים בשכבה הקודמת. CNN שימושי במיוחד לעיבוד תמונה זיהוי תמונות (Image Recognition) וזיהוי אובייקטים בתמונה (Object Detection) כאשר משתני הכניסה מוגדרים על ידי צבעים של עשרות אלפי (או אפילו מיליוני) פיקסלים. רשת נוירונים חוזרת (Recurrent Neural Network) הינה ורסיה של ANN אשר מתאימה במיוחד למצבים שבהם המודל המיועד לניבוי משתנה היציאה צפוי להתפתח על פני זמן.



פרטים אודות כותב המאמר: מדען הנתונים רועי פולניצר, PDS

- מייסד ומנכ"ל האיגוד הישראלי למדעני נתונים מקצועיים (PDSIA), מייסד ויו"ר לשכת מעריכי השווי והאקטוארים הפיננסיים בישראל (IAVFA) ובעלים של פירמת הייעוץ וההדרכה שווי פנימי.
- מחזיק בתואר M.B.A. במנהל עסקים עם התמחות בניהול סיכונים ואקטואריה ותואר B.A. בכלכלה עם התמחות במימון שניהם בהצטיינות מאוניברסיטת בן-גוריון בנגב, דיפלומה בניהול סיכונים פיננסיים (FRM) מאוניברסיטת אריאל, תואר Financial Risk Manage מארגון בינ"ל GARP, תואר Certified Risk Manage מארגון ישראלי IARM, תואר Fellow Actuary מארגון ישראלי IAVFA ותואר Professional Data Scientist מארגון ישראל PDSIA.
- בעל ניסיון אינטנסיבי של מעל עשור וחצי שנים בתחום מדע הנתונים ולמידת המכונה, הכולל ביצוע מחקרי מידע מעמיקים לשם הפקת תובנות עסקיות, ניקוי, טיוב וסידור של המידע המשמש למחקרים השונים, הפעלת אלגוריתמים שונים של מידול, כריית נתונים ו-Machine Learning על המידע ובניית תהליכי הכנת המידע והאופטימיזציה של האלגוריתמים השונים.
- מרצה לתכנות בשפות R ו-Python, לניהול סיכונים, הערכות שווי ואקטואריה והנדסה פיננסית.