|  |
| --- |
|  |
| Plain Deep Neural Network |
|  |
|  |

Creating a Deep Neural Network

**פירוט המשימה:**

**לבנות רשת נוירונים שיודעת להבדיל בין100 קלאסים שוננים של אתרים בעולם.**

**המשימה התחחלקה לשניים:**

**1. ליצור Deep Neural Network שתדע לסווג את התמונות לקלאסים השונים.**

**2. ליצור מסווג KNN שלפי הפיצ’רים של הרשת העמוקה שייצרנו תמצא את K התמונות הקרובות לה ביותר ותציג אותן.**

**פירוט הניסוי:**

**בניסוי זה ננסה ליצור רשת חדשה לגמרי שתדע לטפל בבעית זיהוי האתרים.**

**האתגר של ליצור רשת חדשה גדול בהרבה מהניסוי בו אנו משתמשים ב Transfer Learning מכיוון שכעת נצטרף לקבוע גם את עומק הרשת איזה סוגי שכבות לשים, באיזה סדר וכו’**

**בניוסי זה נשתמש בoptimizer של AdaGrad וב c**

**שלבי ההרצה של ניסוי בודד:**

**- טעינת התמונות של ה TrainingData וטעינת התמונות של ה ValidationData**

**- יצירת המודל (שכבות קונבולוציה + relu, שכבות MaxPooling, שכבות Dense, שכבות Normalization וכו’.**

**- הגדרת פונקציית ה loss ה optimizer וה Hyper Parameters הרלוונטים לאותו optimizer (Batch Size, learning Rate)**

**- יצירת ה callbacks עבור שמירת ההיסטוריה של תוצאות ה loss וה accuracy בסיום כל epoch + שמירת המשקלים.**

**- אימון המודל.**

**- יצירת הגרפים ומדידת תוצאות האימון.**

**חלק ראשון בניית הרשת הבסיסית:**

**בשלב זה נבנה רשת בסיסית שתצליח לסווג בצורה טובה יותר מבחירות רנדומלית (70% Accuracy על ה Validation יהיה מספק מאוד).**

**בשלב זה ניצור בעיקר שכבות קנובולוציה + Relu,שכבות MaxPooling, ושכבות Dense.**

**סיכום תוצאות:**

**חלק שני קביעת המשקלים ההתחלתיים:**

**סיכום תוצאות:**

**חלק שלישי קיבוע ה Hyper Parameters:**

**בשלב זה נקבע את ה Hyper Parameters. בניסוי זה לא נבצע חיפוש מעמיק של ה hyper Parameters האידאלייים (כמו בניסוי שביצענו ב Transfer Learning) מכיון שבניסוי זה נרצה להתמקד בבניית הרשת על כן נבצע שלושה ניסויים רק כדי למצוא ערכים בהן אנו מצליחים לאמן את הרשת בצורה טובה**

**סיכום תוצאות:**

**חלק רביעי הוספת שכבות Normaliztion :**

**סיכום תוצאות:**

**חלק חמישי הוספת שכבות Dropout :**

**סיכום תוצאות:**

**חלק שישי אימון הרשת ושמירת המשקלים שלה:**

**סיכום תוצאות:**

**חלק שביעי – הקטנת ה Learning Rate והגדלת ה Batch Size:**

**כעת הרשת שלנו נמצאת עם למעלה מ 80% accuracy על ה validation שלנו. מה שייגרום לכך שהסיכוי שנעשה צעד בכיוון הלא נכון לגבוה יותר. בשלב זה כדי לאמן את הרשת בצורה טובה יותר נבצע עוד 3 הרצות בהן נקטין בצורה משמעותית את ה Learning Rate שלנו ונגדיל את ה Batch Size.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Epochs | momentum | Batch Size | Learning Rate | Optimizer | Loss Function |
| 10 | 1 | 50 | 0.001 | AdaGrad | categorical\_crossentropy |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Epochs | momentum | Batch Size | Learning Rate | Optimizer | Loss Function |
| 10 | 1 | 80 | 0.0008 | SGD - Momentum | categorical\_crossentropy |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Epochs | momentum | Batch Size | Learning Rate | Optimizer | Loss Function |
| 10 | 1 | 110 | 0.0005 | SGD - Momentum | categorical\_crossentropy |

**סיכום תוצאות:**