

Report

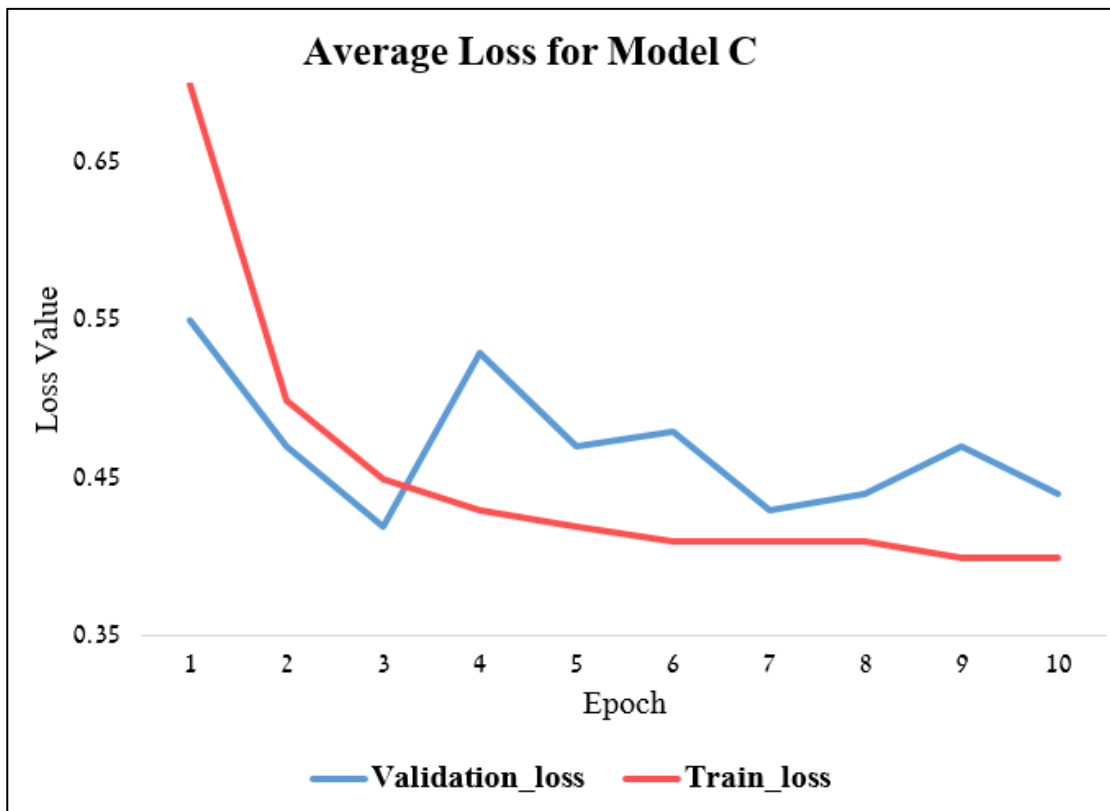
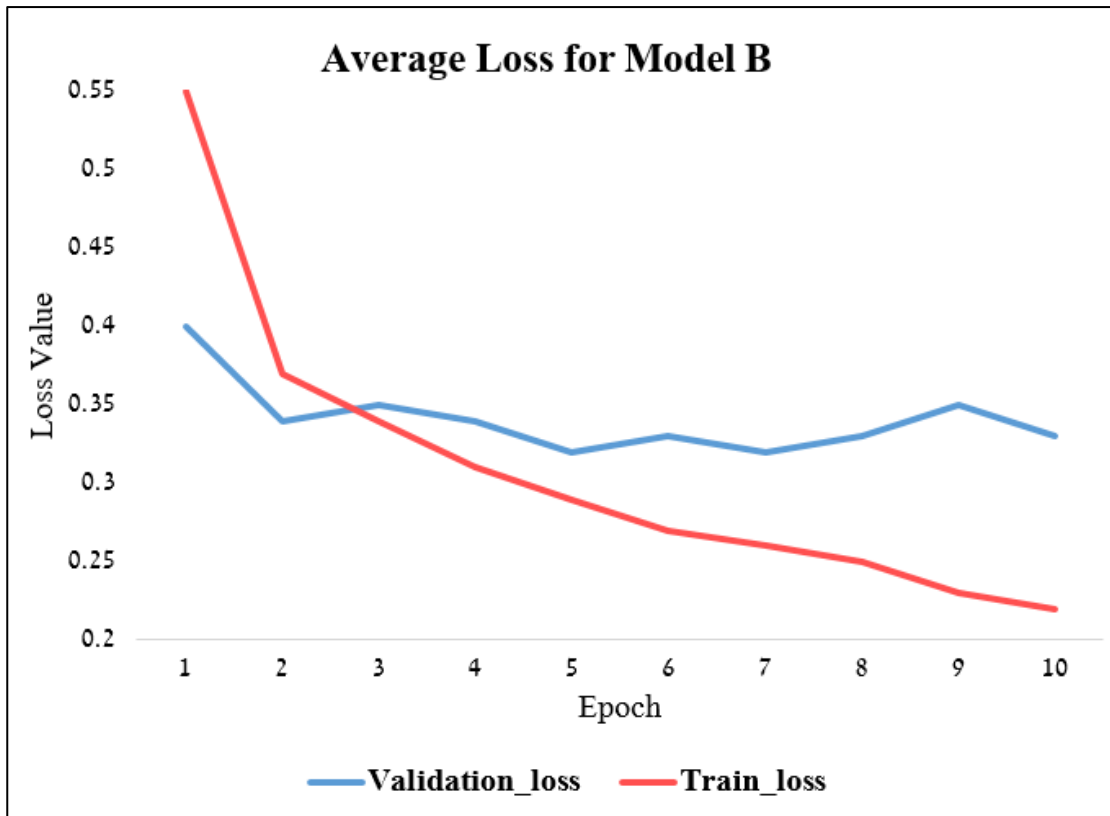
רקע

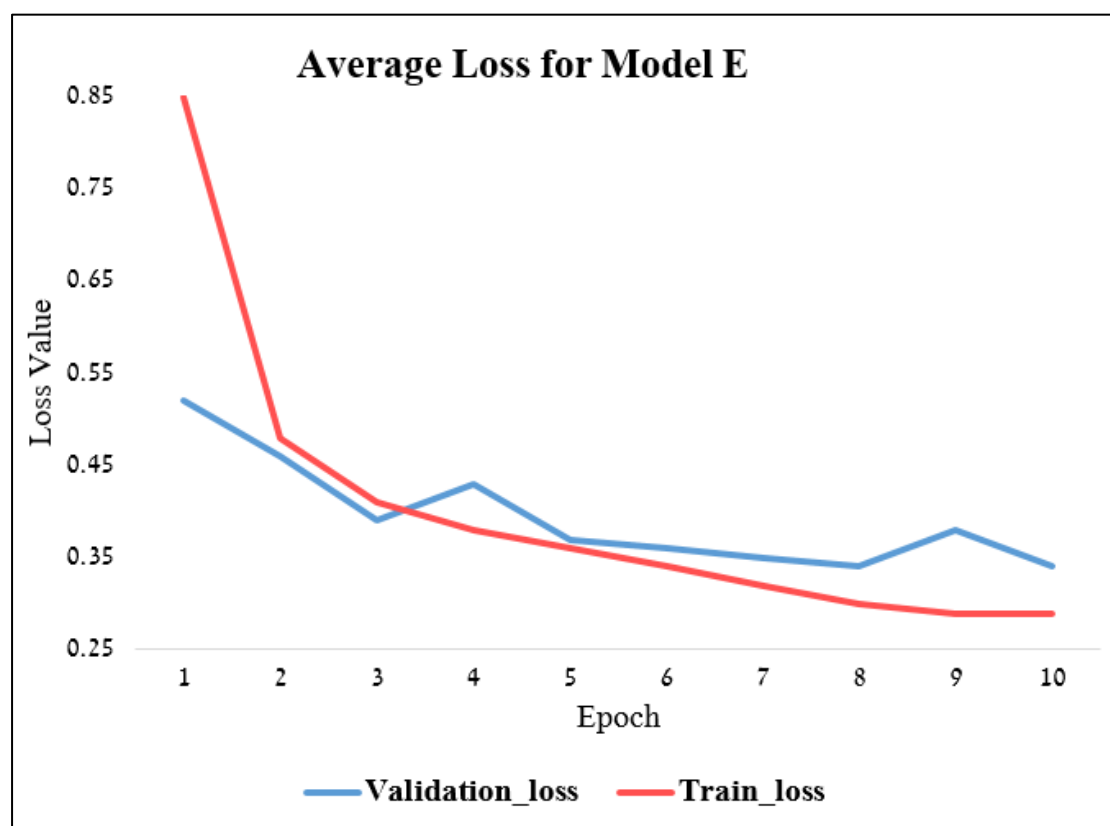
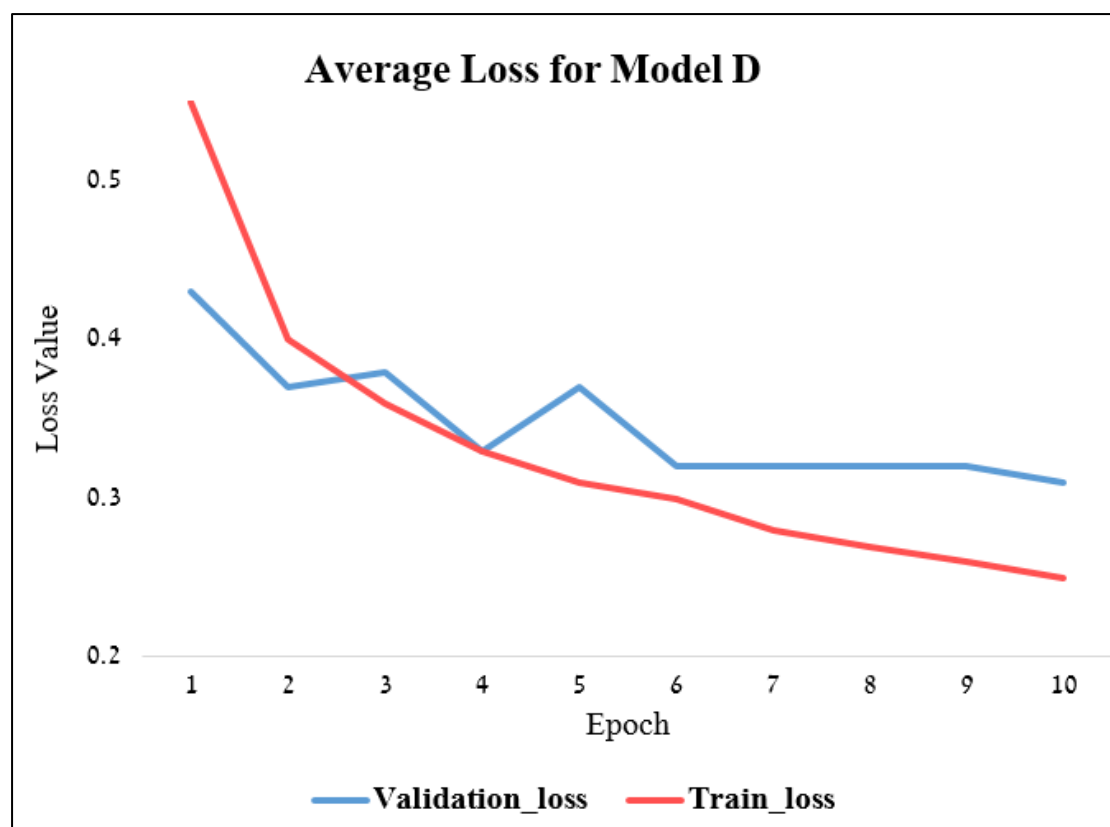
בתרגיל זה התבקשנו לבנות רשת נוירונים שיודעת לסווג כל תמונה לסוג הבגד המופיע בה ע"י שימוש בספריות numpy, scipy וב-PyTorch framework. בסט האימון התקבלו 55,000 תצפיות, כאשר כל תצפית מיוצגת כתמונה של פריט אופנה בעל 28×28 פיקסלים. עבור כל דוגמא בסט האימון קיבלנו את התיוג המתאים לו מבין התיוגים 0,1,2,3,4,5,6,7,8,9, ובסט המבחן התקבלו 5,000 דוגמאות חדשות.

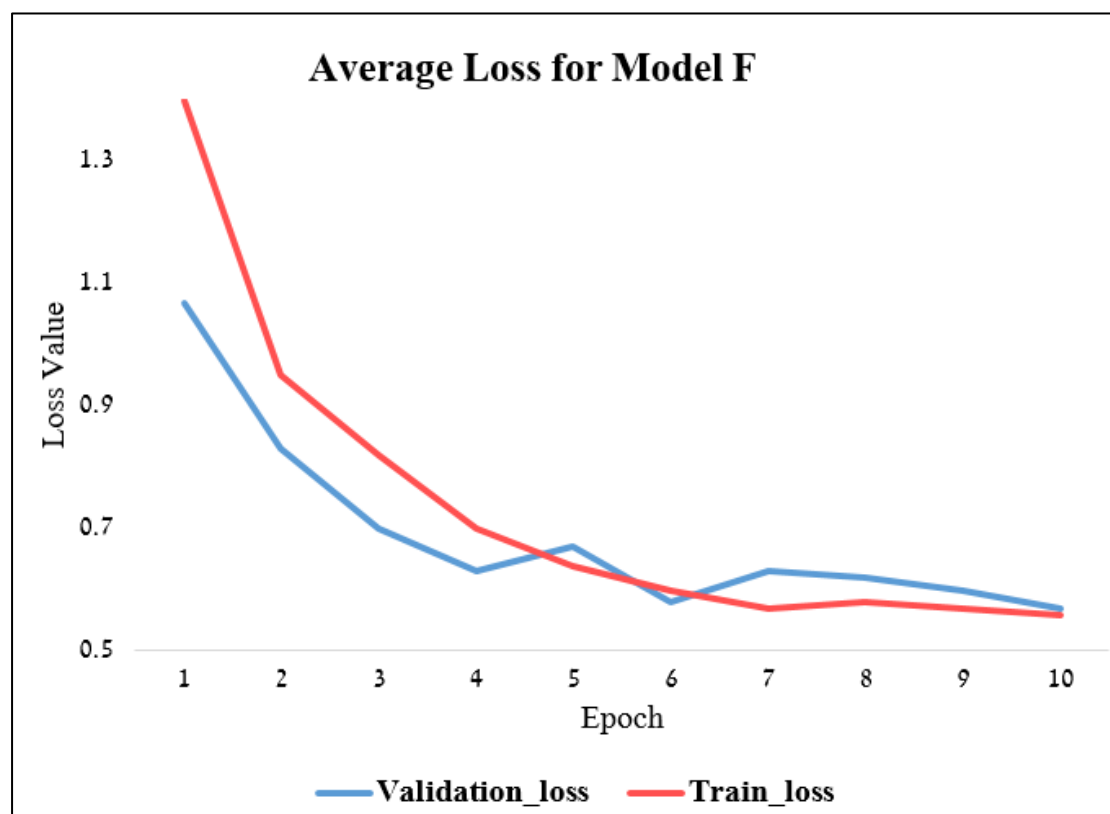
בתרגיל זה בנינו שישה מודלים אפשריים עבור סט הנתונים (Fashion MNIST) שתואר לעיל, ורצינו לבדוק איזה מודל מבניהם ייתן את הביצועים הטובים ביותר.

להלן פונקציות ה-Loss שהתקבלו בשלבי האימון והולידציה:

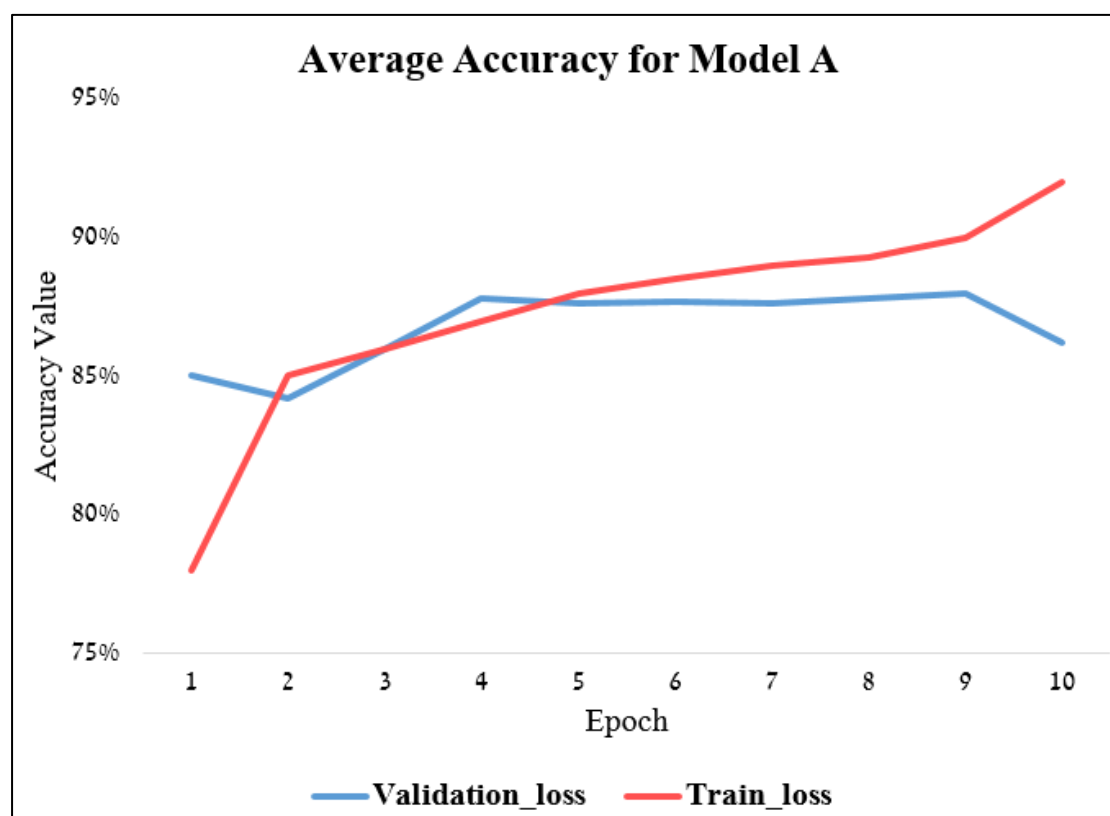


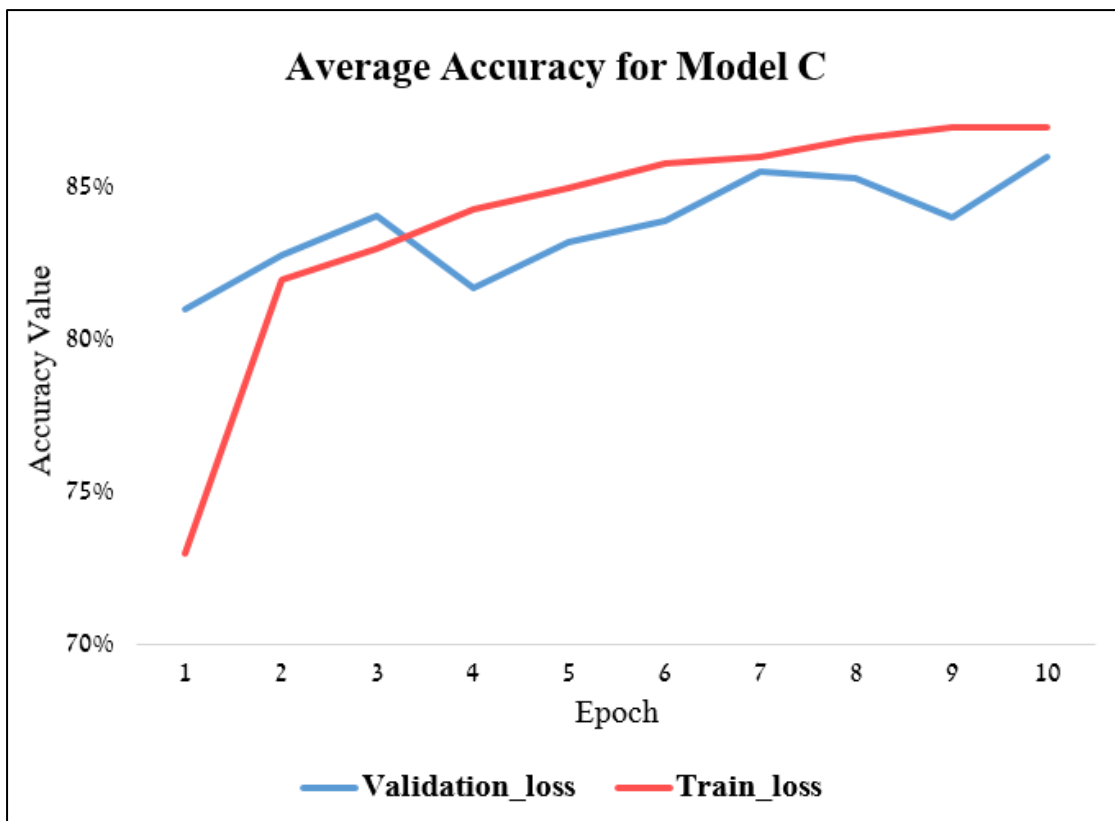
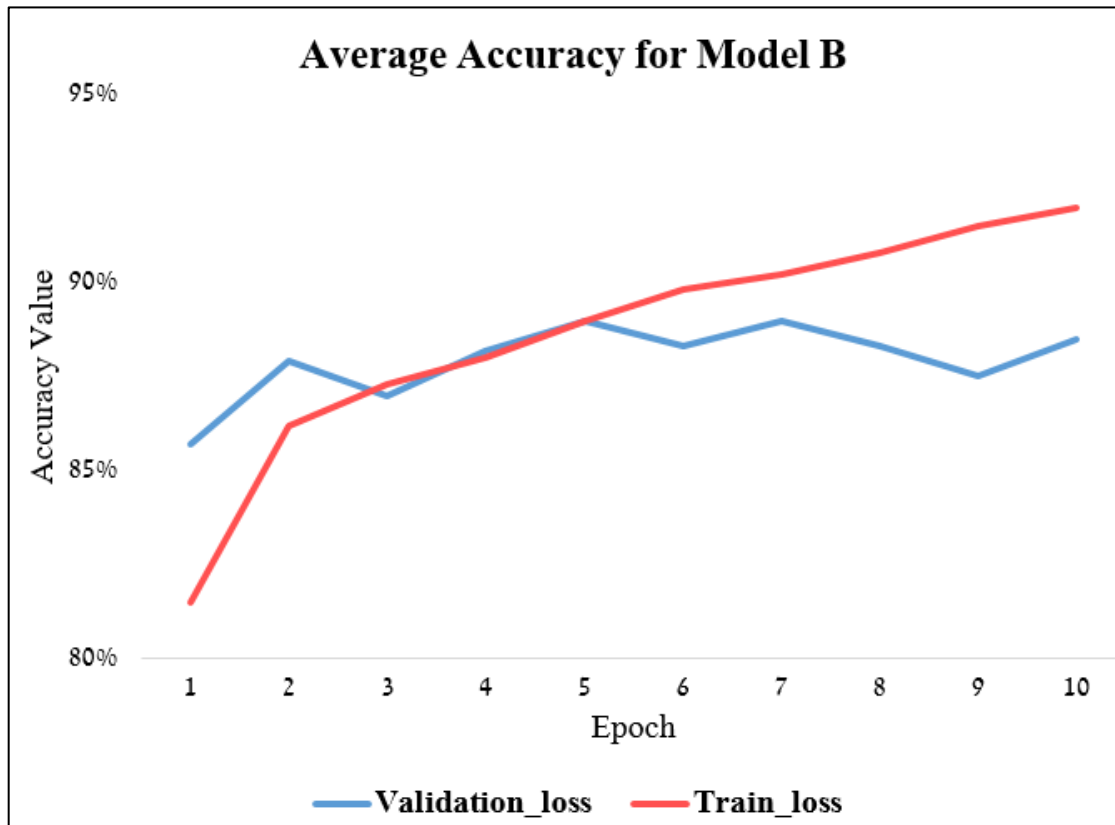


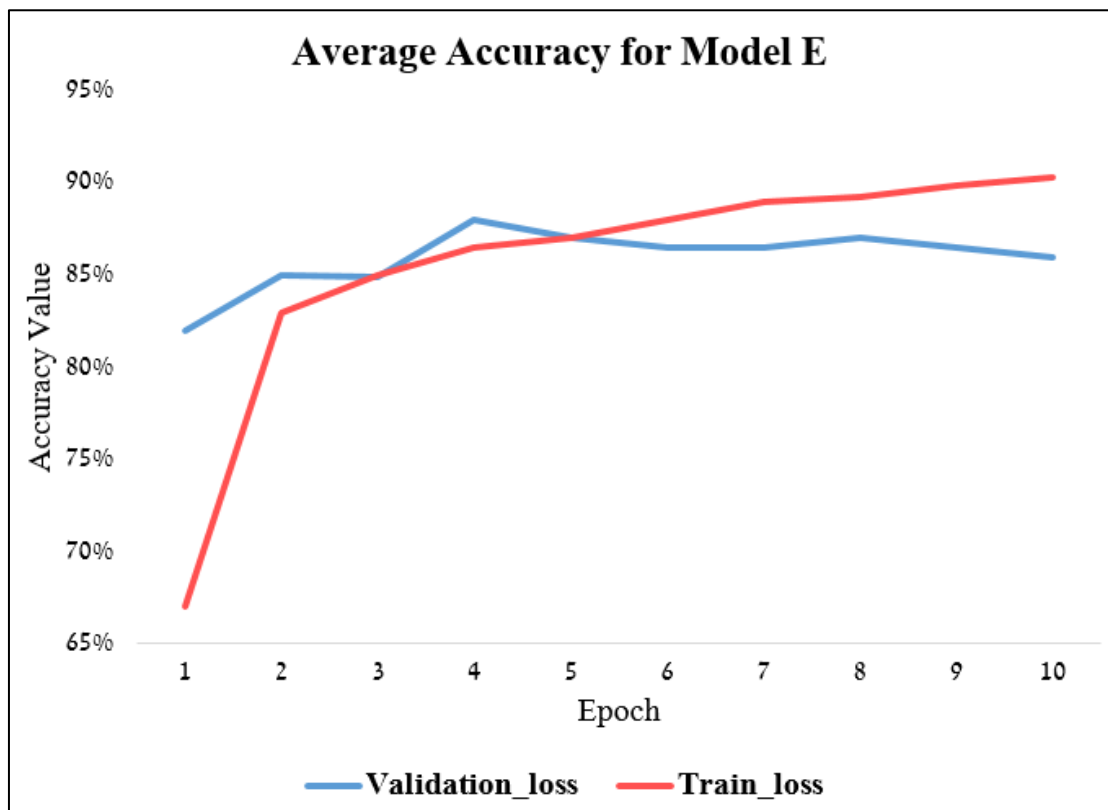
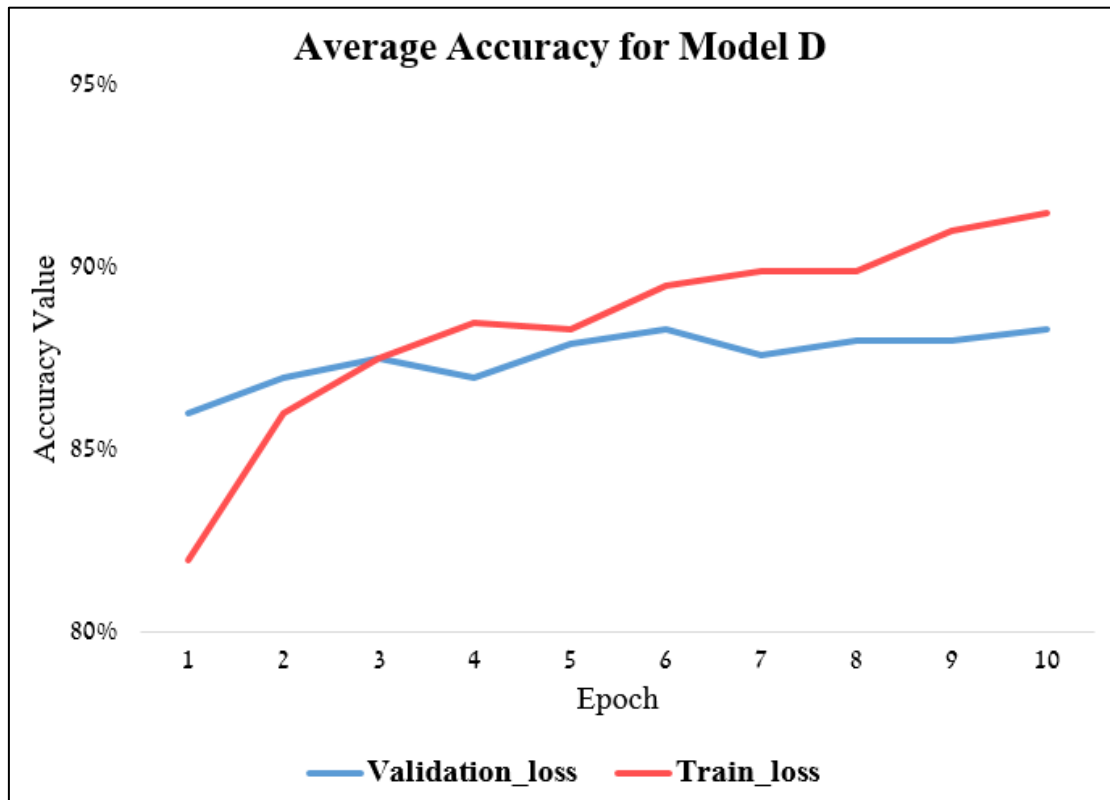


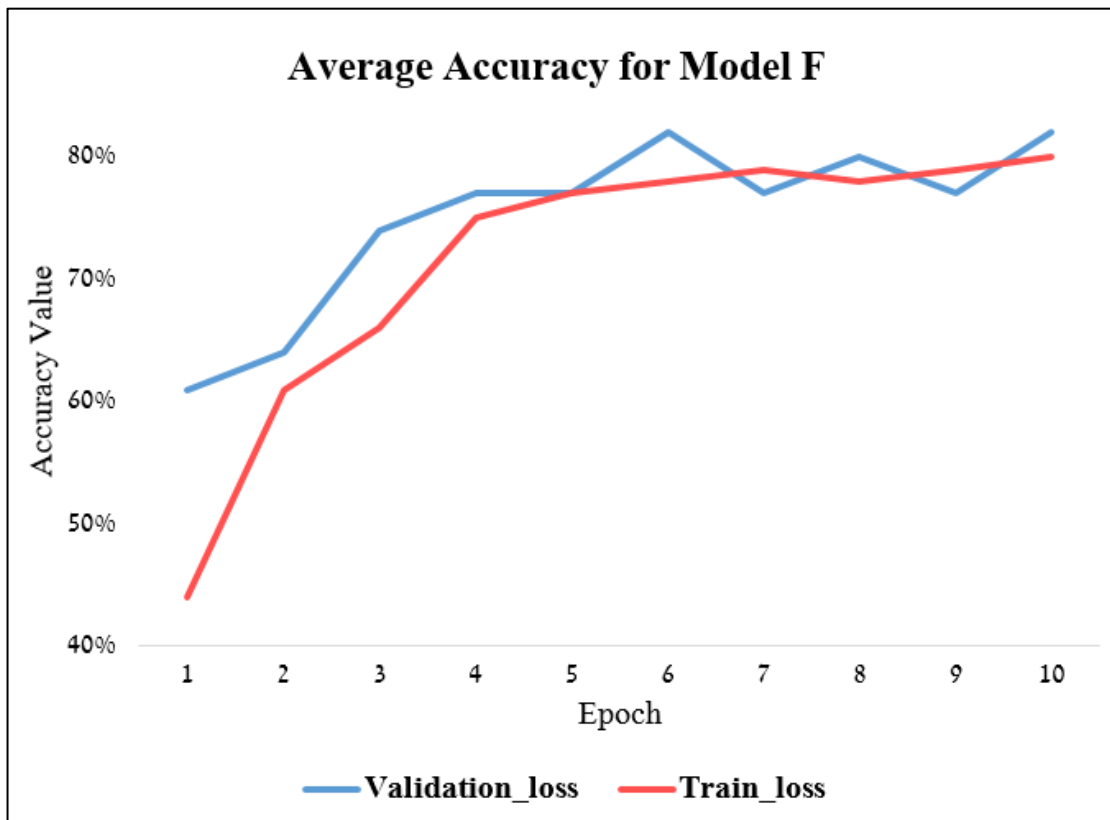


להלן שיעורי הדיוק שהתקבלו בשלבי האימון והולידציה:









שיעורי הדיוק שהתקבלו על סט המבחן המובנה של PyTorch

Model	Test_Accuracy
Model A	87.36%
Model B	87.50%
Model C	85.07%
Model D	88.43%
Model E	87.70%
Model F	83.50%

קביעת היפר פרמטרים ותיאור הרצות המודלים

(1) מספר האפוקים- נקבע להיות 10 עפ"י דרישות התרגיל.

(2) **Learning rate** - ע"י מעבר על הערכים המקובלים [0.1, 0.01, 0.001] עבור כל מודל נבחר ערך קצב הלמידה שנתן אחוזי דיוק יותר גבוהים ופונקציית loss יותר נמוכה. בנוסף עשינו fine tuning לפי הגרפים שקיבלנו על ה- train & validation.

להלן פירוט קצב הלמידה שנבחר עבור כל מודל:

Model	Learning Rate
Model A	0.01
Model B	0.001
Model C	0.1
Model D	0.01
Model E	0.0025
Model F	0.01

(3) **פונקציות האקטיבציה** (ReLU/Sigmoid) נקבעו מראש בדרישות התרגיל עבור כל מודל.

(4) **Optimizers** (SGD/ADAM) נקבעו מראש בדרישות התרגיל עבור כל מודל, פרט למודלים E & F בהם לא נאמר במפורש באיזה להשתמש, ולכן נבדקו במודלים אלו שתי האופציות.

Optimizers שנבחרו למודלים E & F:

Model	Optimizer
Model E	ADAM
Model F	ADAM

(5) **גודל השכבות** עבור הרשתות במודל נקבע מראש בדרישות התרגיל:

Layers_size for Models A to D = [784, 100, 50, 10]

Layers_size for Models E & F = [784, 128, 64, 10, 10, 10, 10]

(6) **batch_size** שנבחרו עבור כל מודל בנפרד:

Model	batch_size
Model A	1
Model B	32
Model C	16
Model D	64
Model E	64
Model F	32

המודל שנבחר

המודל שהציג את הביצועים הטובים ביותר הוא **מודל D** שבו השתמשנו ברשת בעלת 2 שכבות חבויות בגודל [100, 50], פונקציית אקטיבציה ReLU, SGD optimizer, קצב למידה 0.01 ובביטוי Batch Normalization לפני הפעלת כל פונקציית אקטיבציה. מודל זה מציג אחוזי דיוק של 88% ומעלה הן עבור סט המבחן שניתן לנו (קובץ test_x) והן עבור סט המבחן המובנה של PyTorch.