⚡️🚀המאמר היומי של מייק 16.08.24: ⚡️🚀

On the Geometry of Deep Learning

אני ממש אוהב מאמרים שחוקרים מה שקורה בתוך המודלים העמוקים שלנו - הרי לדעתי זה התנאי הכרחי לכך שנוכל להתחיל באמת לסמוך על- AI (לפחות חלקית). ואכן הכותבים מדגישים כי למידה עמוקה, על אף הישגיה המרשימים במגוון תחומים, נשארת עדיין בגדר "קופסה שחורה" עם הבנה חלקית בלבד של אופן פעולתה.

המחברים מנסים להסביר מודלים עמוקים באמצעות ספליינים אפיניים (Affine Splines) שהן למעשה פונקציות רציפות ולינאריות למקוטעין במרחב רב מימד. המחקר מתבונן ברשתות נוירונים מזווית גיאומטרית באמצעות ניתוח של חלוקות הנוצרות על ידי ספליינים אפיניים, המקרבות אותן (הרשתות).

בפרט המחברים דנים בחלוקות של מרחב הקלט לפי הקטגוריות שלו הנוצרות על ידי ייצוג לטנטי (השכבה האחרונה לפני שכבת הסיווג) של הרשת. הבנת החלוקה הזו מסייעת להסביר כיצד רשתות עמוקות לומדות ומייצרות חיזוים עבור קלטים שונים.

המחברים גם דנים במבנים גיאומטריים הנוצרים על ידי משקלי המודל במרחב הלוס (כלומר מנתחים את פונקציית הלוס למשקלי הרשת השונים). בנוסף המאמר גם מדבר על החלוקות הנוצרות במרחב משקולות המודל בשכבות שונות לאתחולי רשת שונים וגם לאימון עם ובלי BatchNorm. כמובן שזה נעשה על דוגמאות מלאכותיות(toy examples) בעלי מימד נמוך. ויש עוד מספר ניתוחים גיאומטרים די מעניינים במאמר.

מעניין כי המחברים כותבים כי אחת המטרות המרכזיות של המחקר היא לדרבן מתמטיקאים לעסוק בניתוח גיאומטרי של רשתות עמוקות.

https://arxiv.org/abs/2408.04809