

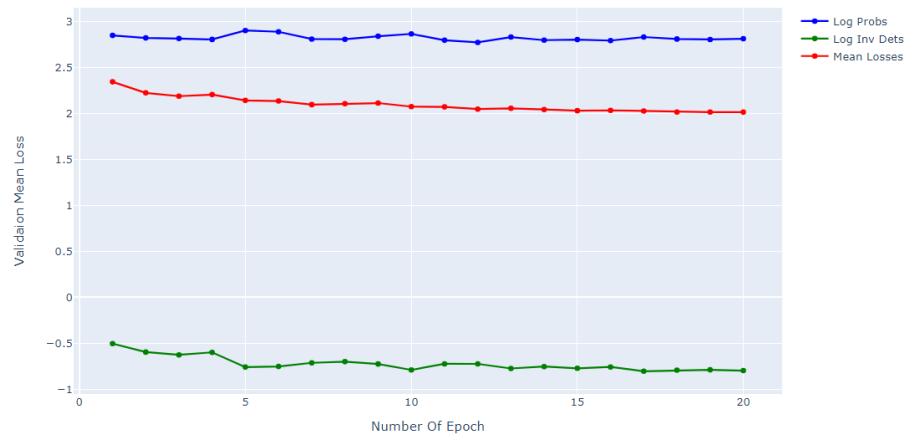
קורס מתקדם בלמידת מכונה – תרגיל בית 2

בר רוז, 203765698

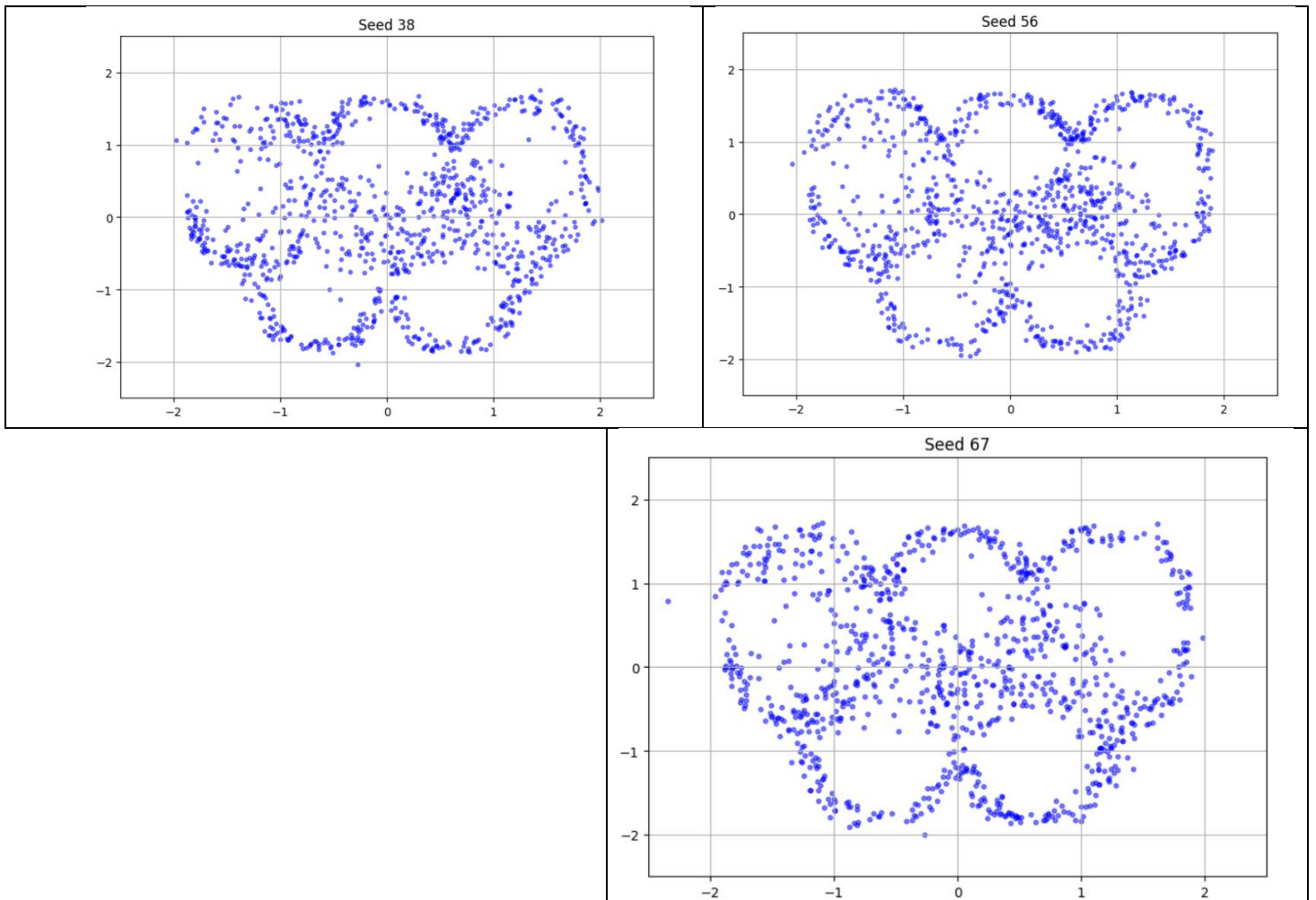
סעיף 3.3 (normalization flow):

שאלה 1:

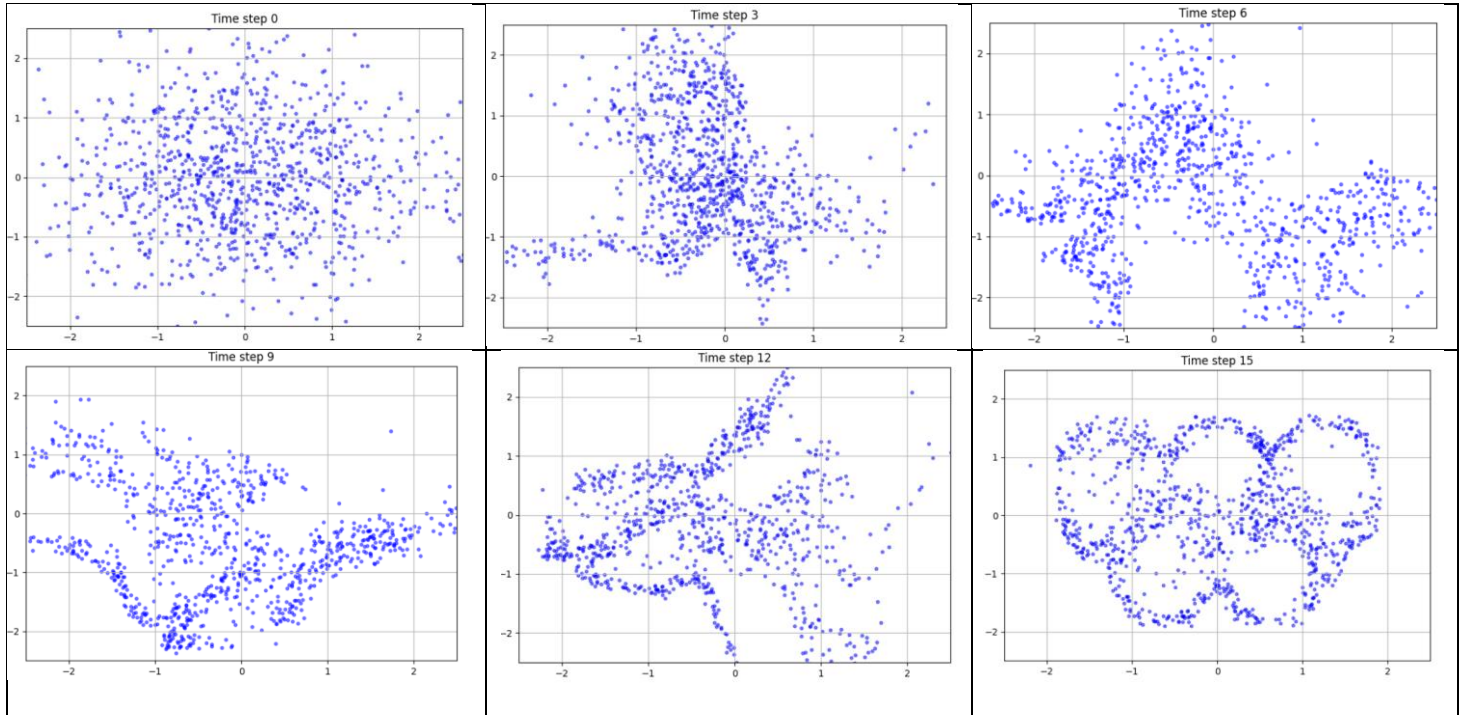
Validation mean losses as functions of number of epoch



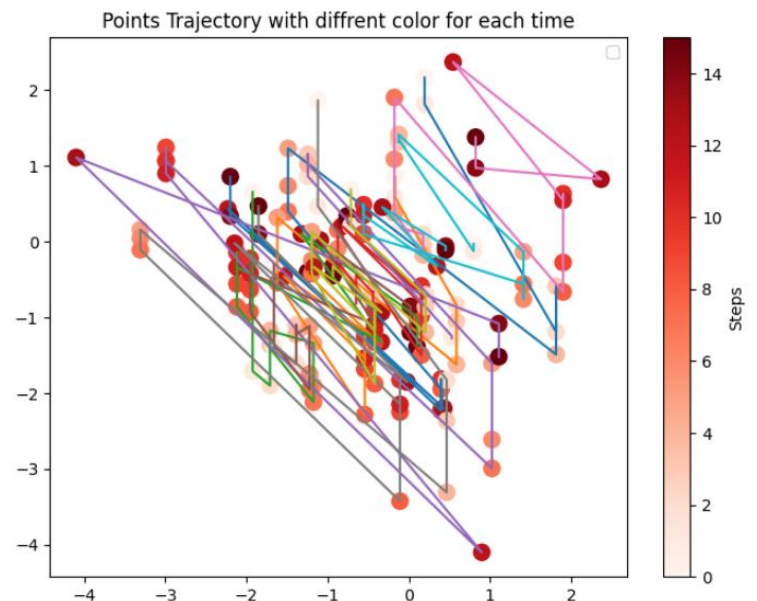
שאלה 2:



שאלה 3:



שאלה 4:



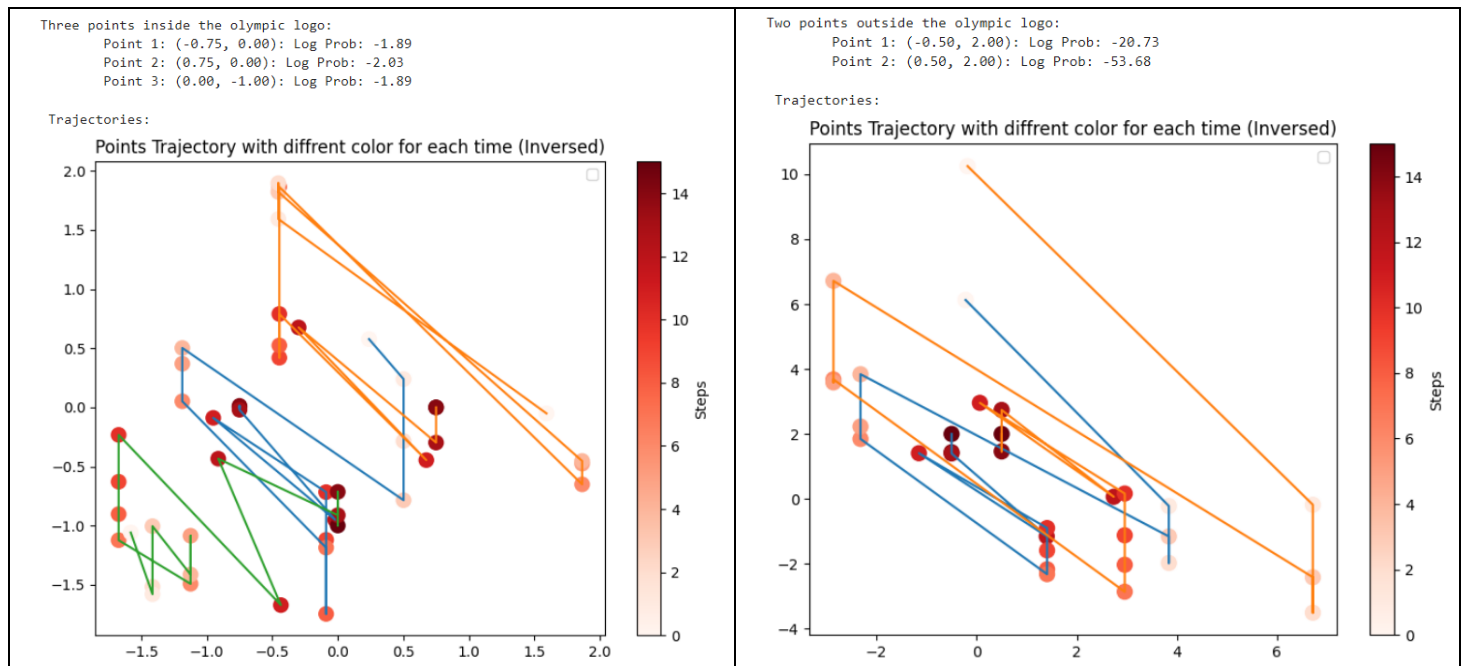
הסבר על הגרף:

צבעו של כל קו מסמל מסלול של נקודה שונה.

צבע של כל נקודה מסמן את המיקום של הנקודה הספציפית בזמן / צעד הספציפי:

צבע לבן - הצעד / זמן ה- 0 – כלומר המיקום של הנקודה כאשר היא נדגמה מהתפלגות גאוסיאנית סטנדרטית.

צבע הכי כהה – הצעד / זמן האחרון – כלומר המיקום של הנקודה בהתפלגות שאנחנו רוצים לחזות (לוגו האולימפיאדה)



הערה: ההסבר על הגרפים הוא בדיוק כמו בשאלה 4.

קיבלנו כי הלוג הסתברות של שלושת הנקודות שנלקחו מתוך ההתפלגות של הלוגו גבוהים לעומת הלוג הסתברות של שתי הנקודות שנלקחו מחוץ להתפלגות של הלוגו.

ניתן לאמת זאת על ידי הגרפים עצמם – אנחנו רואים שהנקודות שנדגמו מתוך הלוגו מתחילות (צבע לבן) בערכים יותר קרובים לתוחלת של ההתפלגות הנורמלית הסטנדרטית, זאת לעומת הנקודות שנדגמו מחוץ ללוגו שמסתיימות בערכים יותר רחוקים מהתוחלת של ההתפלגות הנורמלית הסטנדרטית. לכן מהנוסחה שראינו (למטה) בכיתה נוסע כי תהיה הסתברות גבוהה יותר לנקודות שנדגמו מתוך ההתפלגות של הלוגו.

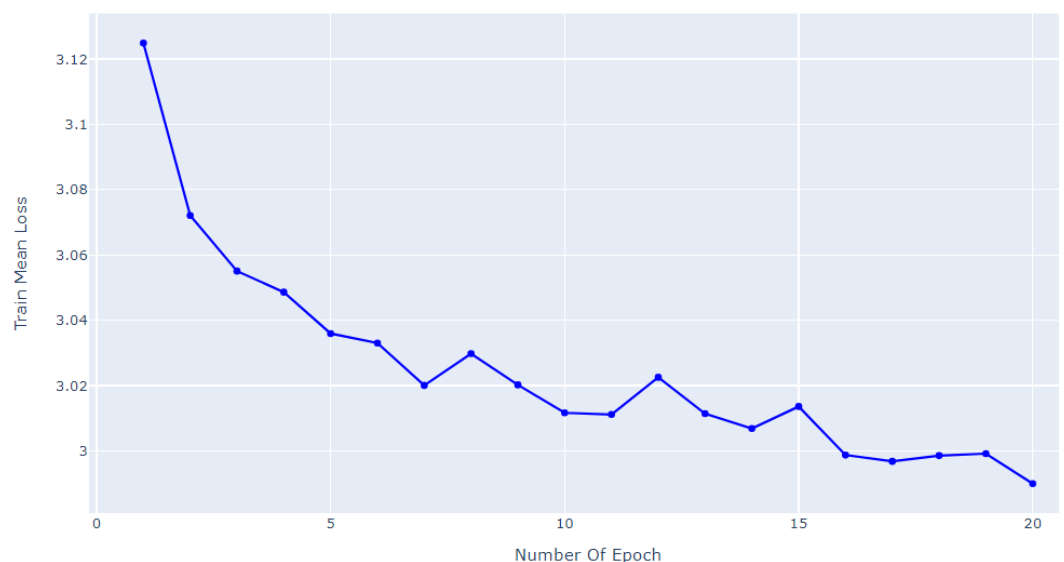
$$p(y) = p_Z(\phi^{-1}(y)) \left| \det \frac{\delta \phi^{-1}(y)}{\delta y^T} \right|$$

סעיף 4.3 (flow matching):

Unconditional Flow Matching

שאלה 1:

Train mean losses as functions of number of epoch

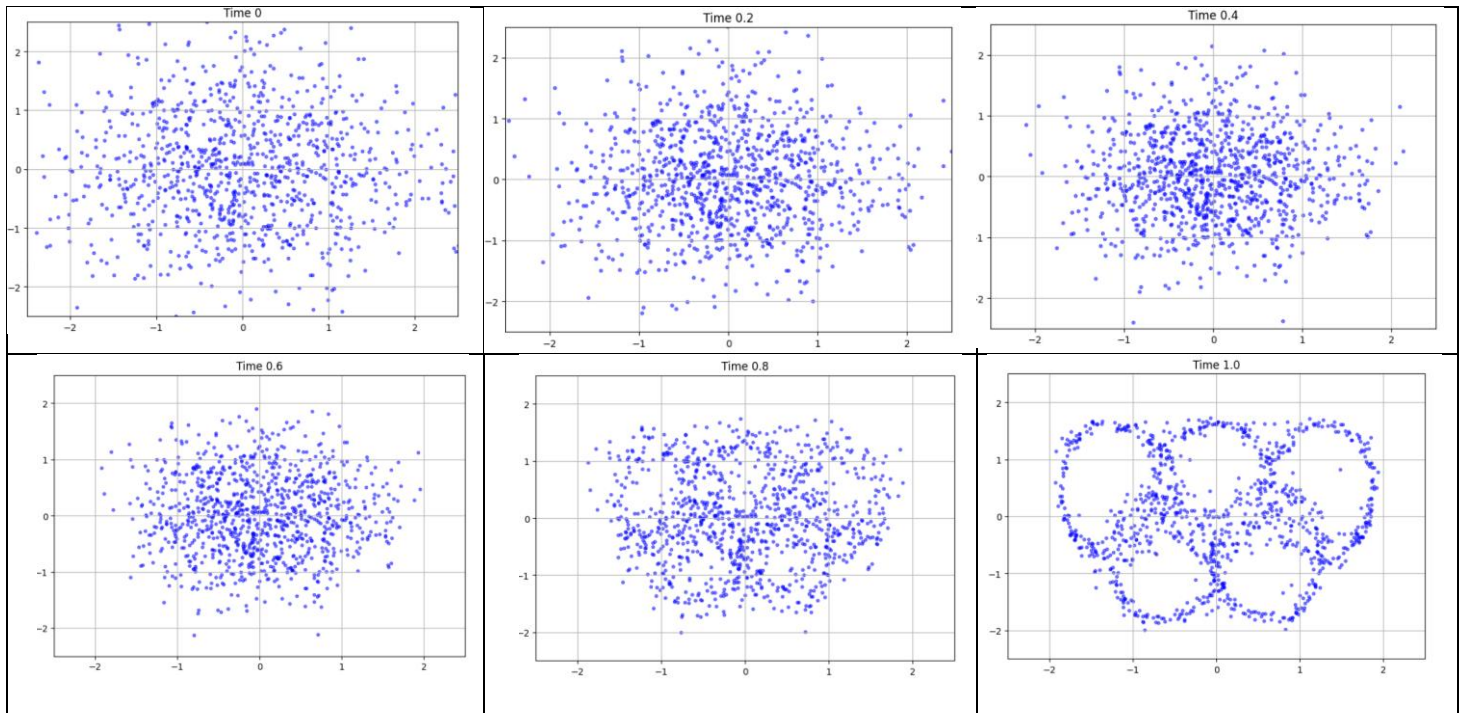


היערה: אומנם ניתן לראות שהלוס הולך וקטן ככל שמספר האפוק גדל, אך בהשוואה לגרף לוס שאנשים אחרים קיבלו אני רואה שהם קיבלו ערכים יותר נמוכים (אצלי מתכנס ל3 בעוד שאצלם מתכנס ל1.5).

לצערי, אחרי **שעות ארוכות** של דיבאג לא הצלחתי למצוא את הבעיה.

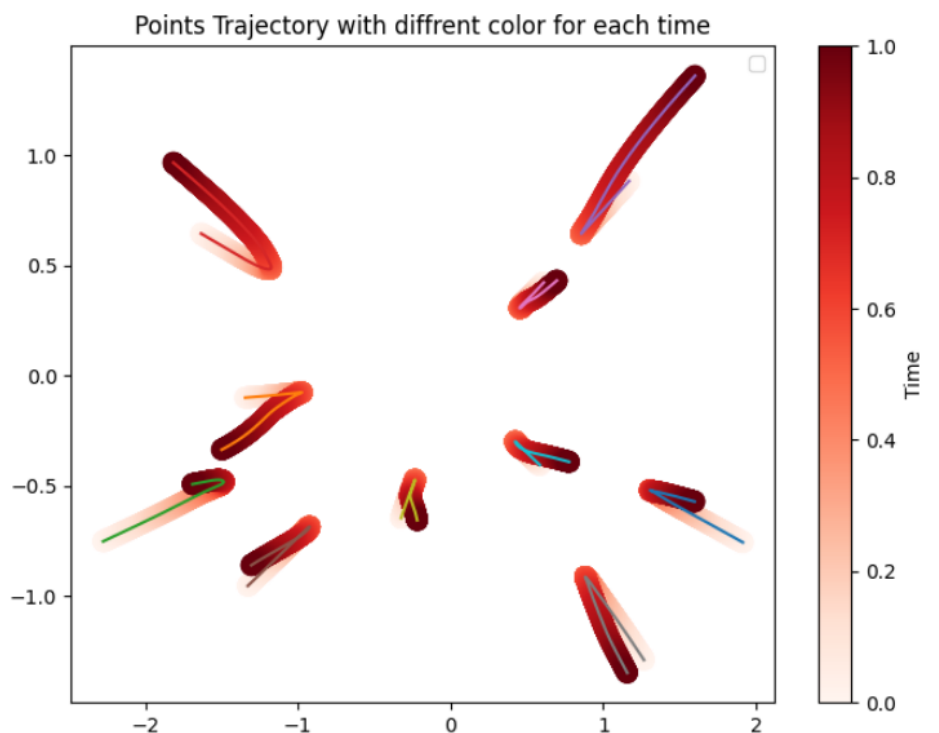
אף על פי כן, נראה שההתפלגות המתקבלת עדיין נראית טובה ביחד למצופה.

שאלה 2:



normalization flowa בכל צעד בחרנו במימד אחד של הווקטור ושינינו את הערך שלו בעוד שהערך במימד השני לא השתנה כלל. לעומת זאת בflow matching, כל נקודה השתנתה בזמן בהתאם לשדה הווקטורי שחישבנו. מסיבות אלה נובע שבכל צעד בnormalization flowa התוחלת של כל הנקודות משתנת בהתאם לציר שעליו אנחנו עובדים בעוד שבflow matching התוחלת של הנקודות נשארת קבועה (אפס). ואכן ניתן לראות זאת בגרפים.

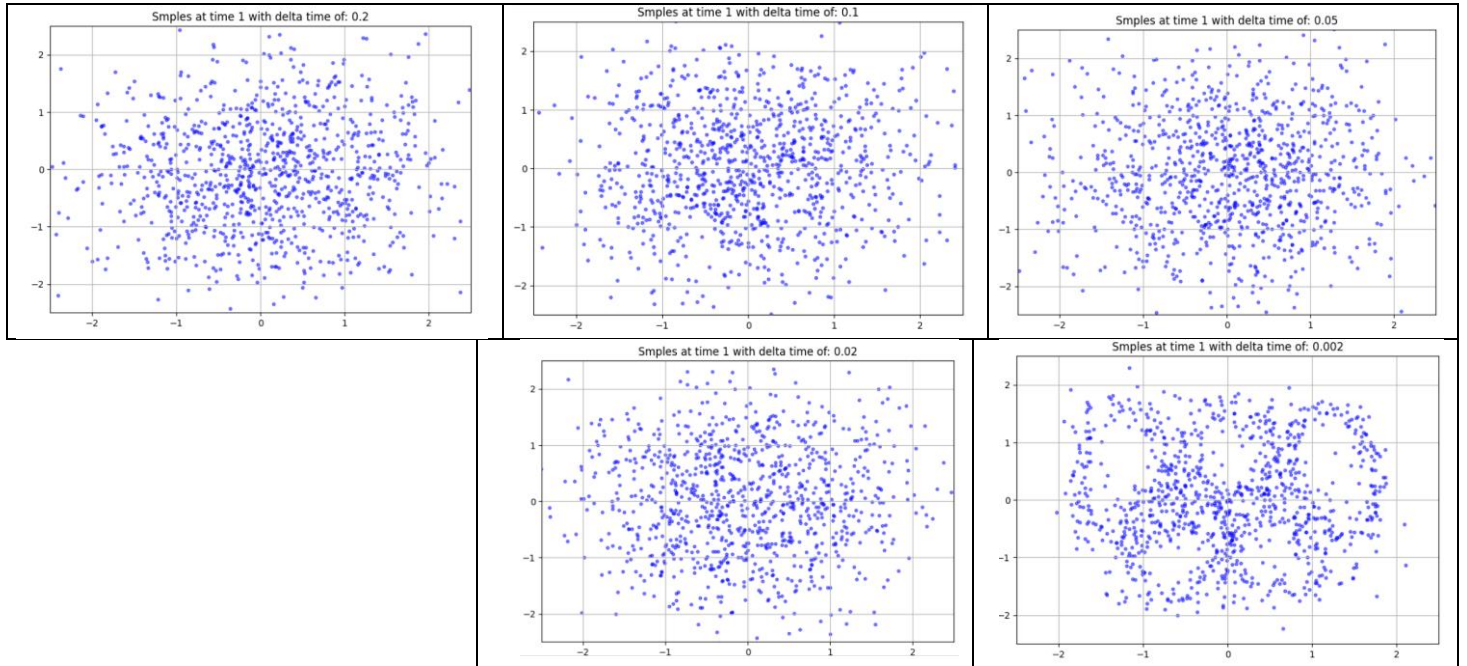
שאלה 3:



היערה: ההסבר על הגרפים הוא בדיוק כמו בשאלה 4 של normalization flowa.

ניתן לראות כי לעומת הגרף בשאלה 4, קיבלנו שכל נקודה משתנה הרבה יותר לאט בסה"כ. ניתן להסביר זאת בדומה לסעיף הקודם, כאשר בnormalization flow אנחנו עובדים בכל צעד על ציר אחר בנפרד מה שגורם לשינויים גדולים במיקום של הנקודה, זאת לעומת flow matching בהם השינויים במסלול של כל נקודה הרבה יותר עדינים.

שאלה 4:



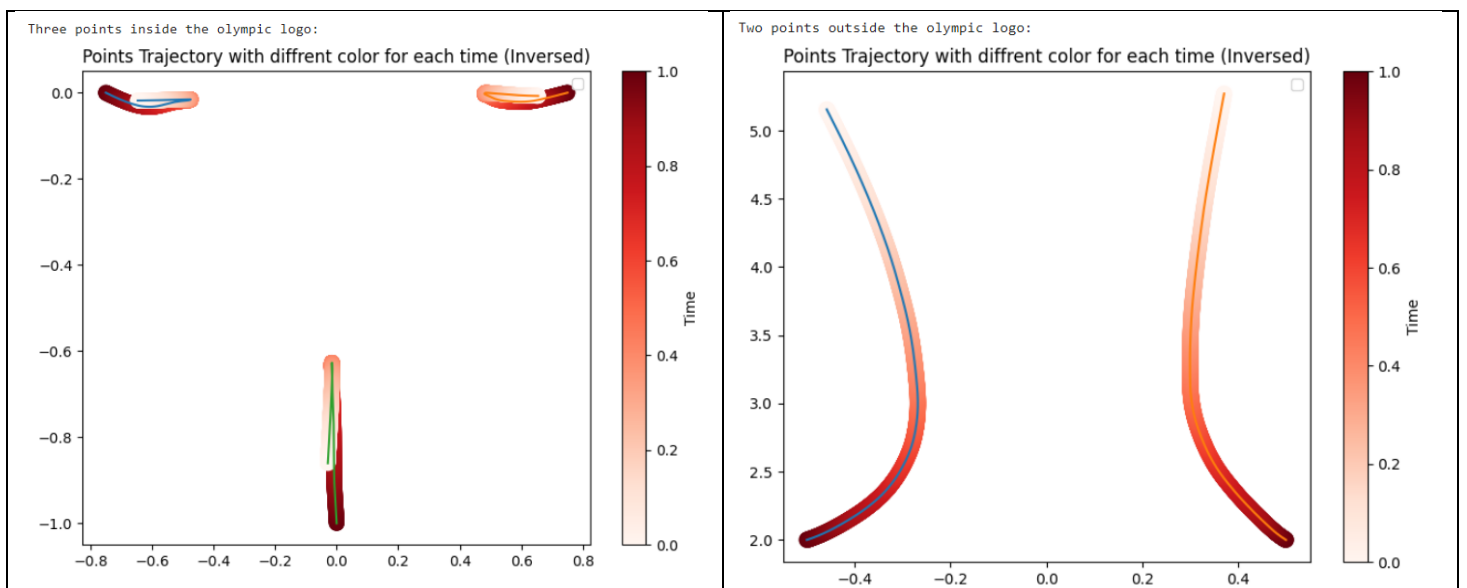
באופן כללי כאשר אנחנו מחשבים forward של flow matching ע"י הנוסחה:

$$y_{t+\Delta t} = y_t + v_t(y_t)\Delta t$$

אנחנו למעשה מבצעים מחשבים את המיקום של כל נקודה בזמן הבא על ידי קירוב טיילור מסדר ראשון. מכאן, נוכל לצפות כי ככל שנשתמש בדלתא זמן יותר קטן, כך נתכנס להתפלגות המטרה (הלוגו) בצורה יותר טובה.

היערה: יתכן שקיבלתי בסעיף זה תוצאות פחות טובות לעומת אנשים אחרים ומתחילים לראות את הלוגו בערכי דלתא יותר גדולים. זה ככל הנראה נובע מהבעיה אותה תיארתי בסעיף 1.

שאלה 5:

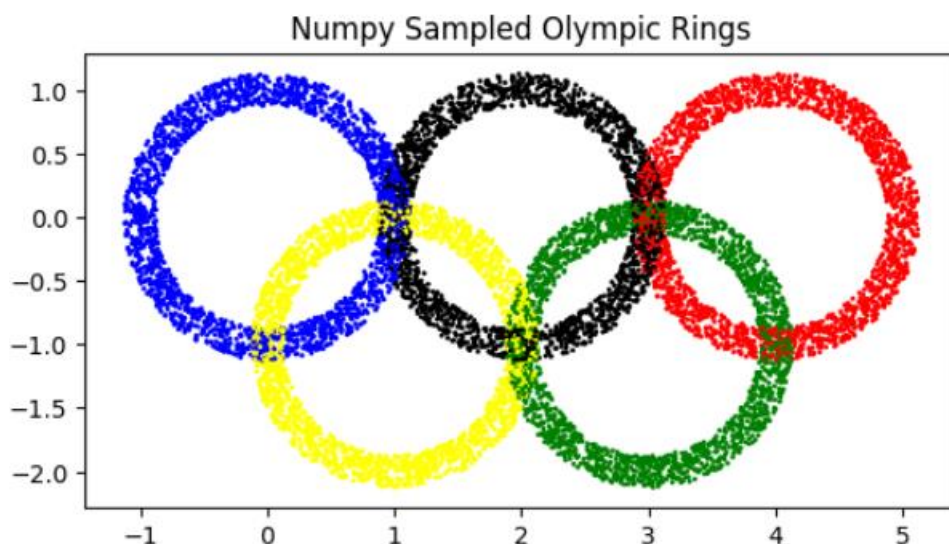


בדומה לגרפים בnormalization flow אנחנו מקבלים שנקודות שנדגמו מחוץ להתפלגות הרצויה (הלוגו) מתחילות מנקודות יחסית רחוקות מהתוחלת של התפלגות הגאוסיאנית הסטנדרטית (ווקטור האפס), זאת לעומת נקודות שנדגמו מתוך ההתפלגות הרצויה (הלוגו) מתחילות מנקודות יחסית קרובות לחולת של התפלגות הגאוסיאנית הסטנדרטית. כמו כן, בדומה לסעיף 3 (מחלק זה) אנחנו רואים שהשינויים במסלול של כל דגימה יותר קטנים בflow matching לעומת normalization flow. מאותו הסבר שהסברנו.

Conditional Flow Matching:

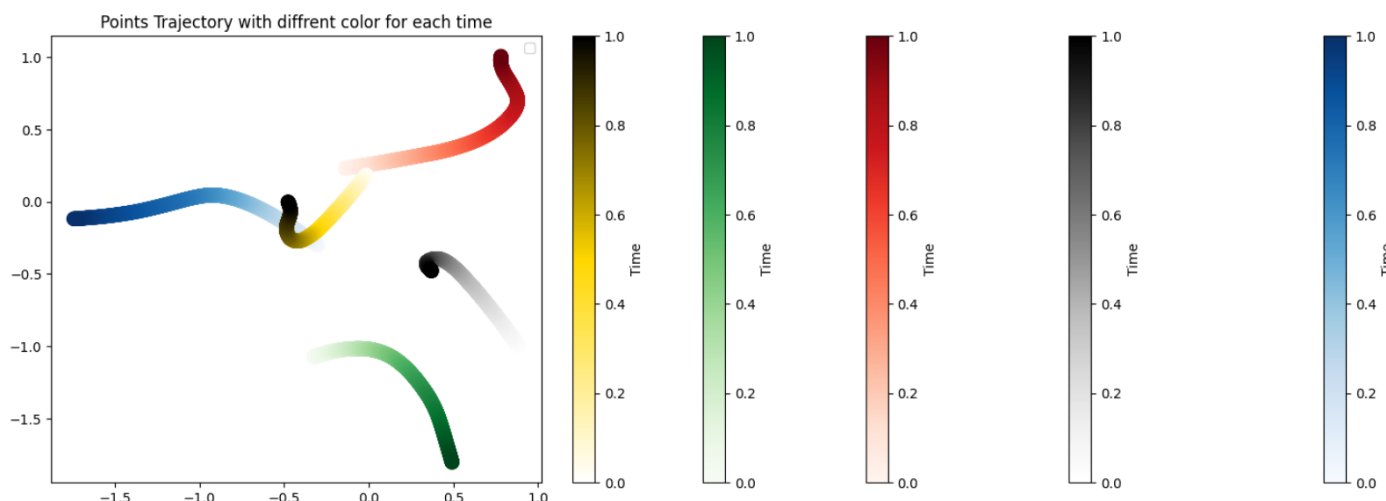
שאלה 1:

הנקודות שאנחנו מכניסים בתור input למודל כדי לאמן אותו:



על מנת להכניס את הconditioning, שרשרתי לאת הembeddings vectors שקיבלנו עבר קלאס (צבע) של כל נקודה לנקודה עצמה ויחד עם הזמן הרלוונטי נתתי למודל שקובע את השדה הווקטורי לחשב את הנגזרת כתלות בזמן.

שאלה 2:



בדומה לסעיפים קודמים, ככל שהצבע יתור בהיר אנחנו נמצאים בזמן יותר מוקדם וככל שהצבע יותר כה אנחנו נמצאים בזמן יותר מאוחר. אכן, לפי הגרף שקיבלנו בסעיף 3 (למטה) ניתן לראות שמסלול של כל נקודה מסתיים בטבעת מאותו הצבע.

