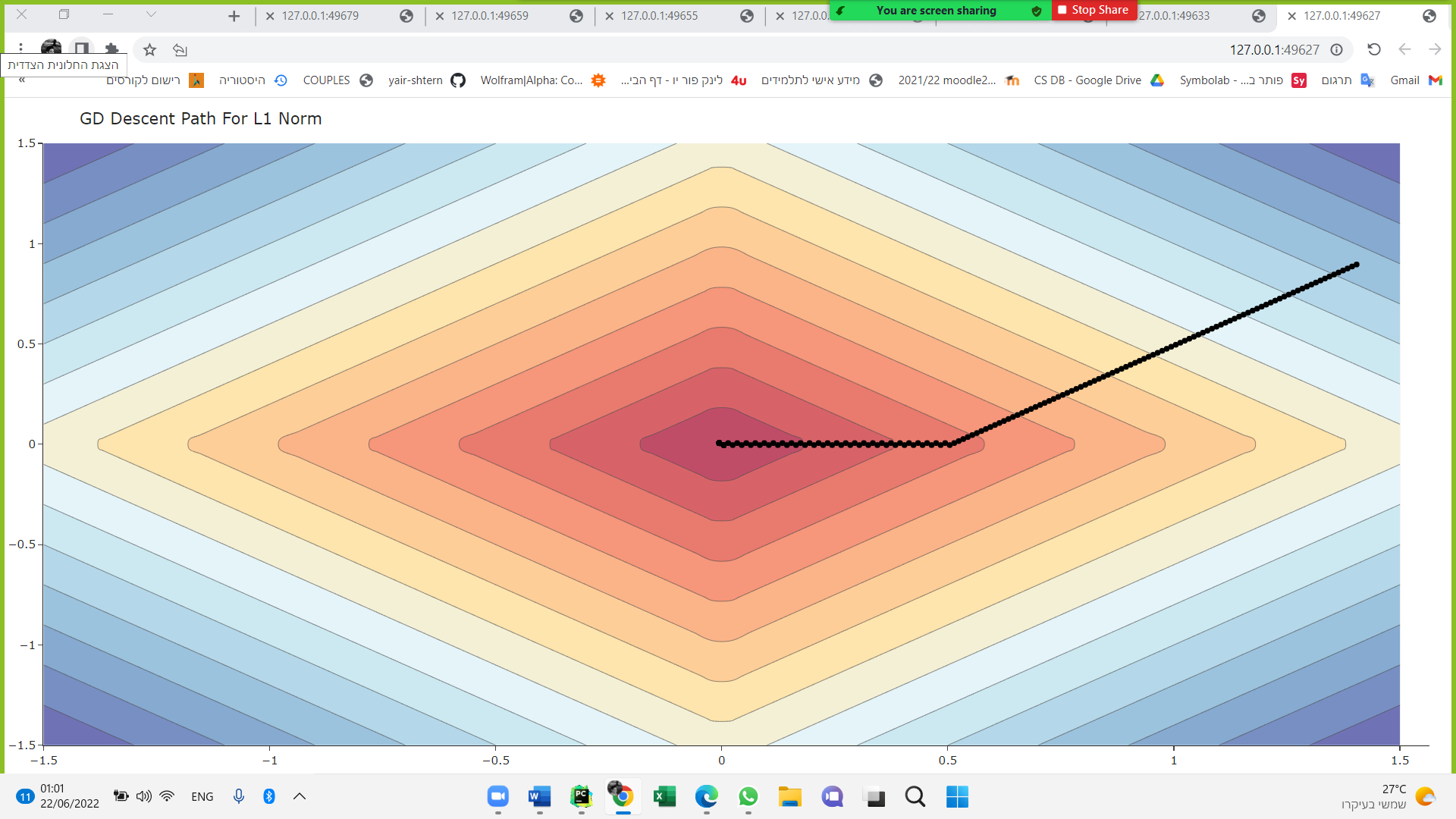
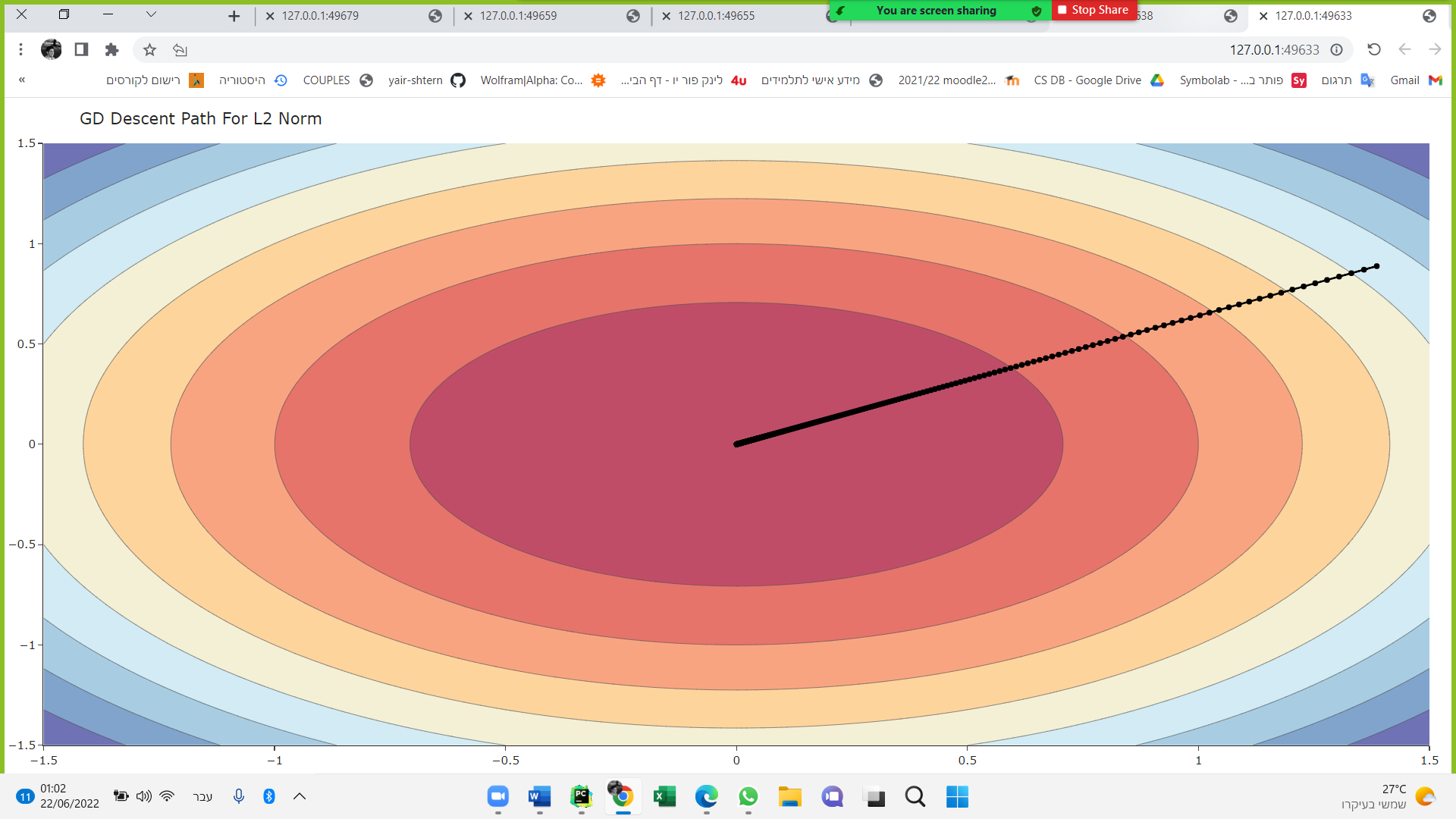
שאלה 1:

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

נשים לב כי היעקוביאן של L2 קטן בצורה משמעותית יותר ככל שהאלגוריתם מתקרב למינימום ולכן גודל הצעד על הפונקציה קטן בכל איטרציה וזה עקב השיפוע התלול של נורמת L2. לעומת זאת, ב L1 השיפוע קבוע וזה גורם לכך שהצעד על הפונקציה ישאר קבוע בכל איטרציה (כי η קבוע ). נשים לב כי עקב גודל הצעדים האלגוריתם עובר לחפש את המינימום גם כאשר L1 נמצאת בעלייה (שכן הוא מרחיק יותר מידי) מה שגורם לשינוי כיוון של הגרדיאנט, בעוד שב L2 כיוון הגרדיאנט נשאר קבוע ורק הגודל שלו משתנה.





שאלה 2:

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

נבחין שישנה נקודה במסלול על נורמת L1 בה המסלול משנה כיוון כלפי מעלה, יורד למטה וכו' כך שהקו מתקרב למינימום אך לבסוף מתבדר ואינו מתכנס אליו, ולכן מינימום לא יתקבל לכל מספר של איטרציות. וזה כי קצב הלמידה קבוע ולכן אם האלגוריתם לא נפל בדיוק בנקודה בה הכיוון השתנה הוא ימשיך לחפש אותו בכיוון ההפוך בכל איטרציה מחדש, כך שהגרדיאנט ישנה את הכיוון כל איטרציה.

כמו כן גודל הצעדים נשאר קבוע שכן השיפוע קבוע ו־η קבועה לאורך כל התקדמות האלגוריתם

שאלה 3:

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

ניתן לראות את חשיבות הבחירה של η. עבור η גדולה מידי נקבל שהאלגוריתם לא מצליח להגיע למינימום ונתקע באותם הערכים במספר האיטרציות הנתון. עבור η לא גדולה מידי ולא קטנה האלגוריתם מצליח להגיע לערך הנמוך ביותר במספר האיטרציות הנתון. עבור η קטנה מידי האלגוריתם מתקדם בצעדים קטנים מידי ולא מצליח להגיע למספר קרוב למינימום במספר האיטרציות הנתון.



שאלה 4:

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

נשים לב כי אכן ה־ loss הקטן ביותר מתקבל עבור 01.0 η = בשתי הנורמות.

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

שאלה 5:

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

ניתן לראות את חשיבות הבחירה של קצב הדעיכה. עבור γ גדולה מידי נקבל כי האלגוריתם לא מצליח להגיע למינימום ונתקע באותם הערכים במספר האיטרציות הנתון. עבור γ לא גדולה מידי האלגוריתם מצליח להגיע לערך הנמוך יותר במספר האיטרציות הנתון. עבור γ קטנה מידי האלגוריתם מתקדם בצעדים קטנים מידי ולא מצליח למספר קרוב למינימום במספר האיטרציות הנתון. תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

שאלה 6:

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

נשים לב כי הפעם ה־ loss המינימלי אשר התקבל עבור L1 כאשר קצב הלמידה אקספוננציאלי קטן משמעותית מאשר ה־loss המינימלי שהתקבל עבור קצב הלמידה הקבוע ההבדל נובע מכך שקצב הלמידה האקספוננציאלי מתאים את קצב הלמידה עם ההתקדמות ולכן פחות תלוי בנקודת ההתחלה של האלגוריתם.

תמונה שמכילה טקסט

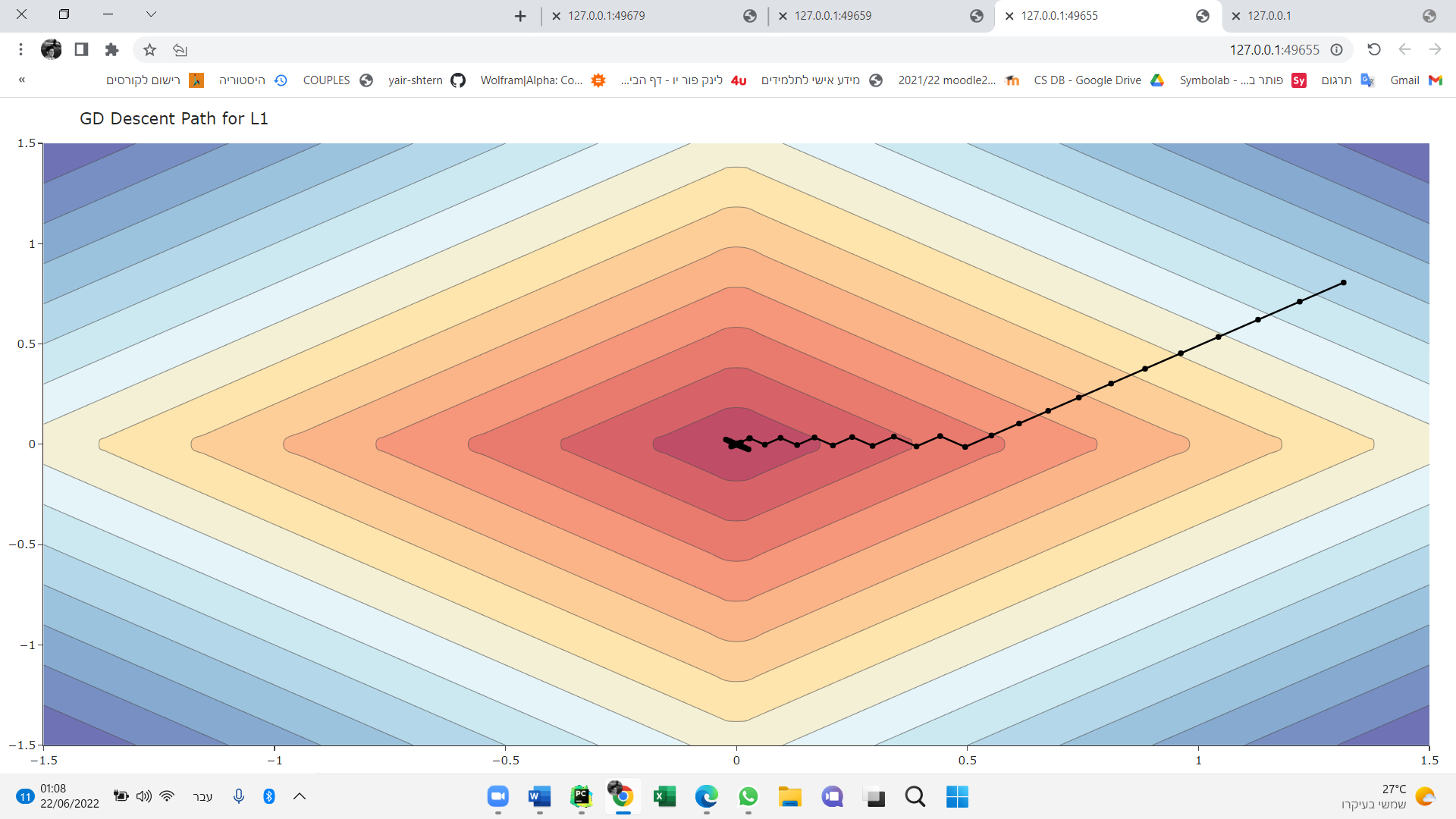
התיאור נוצר באופן אוטומטי

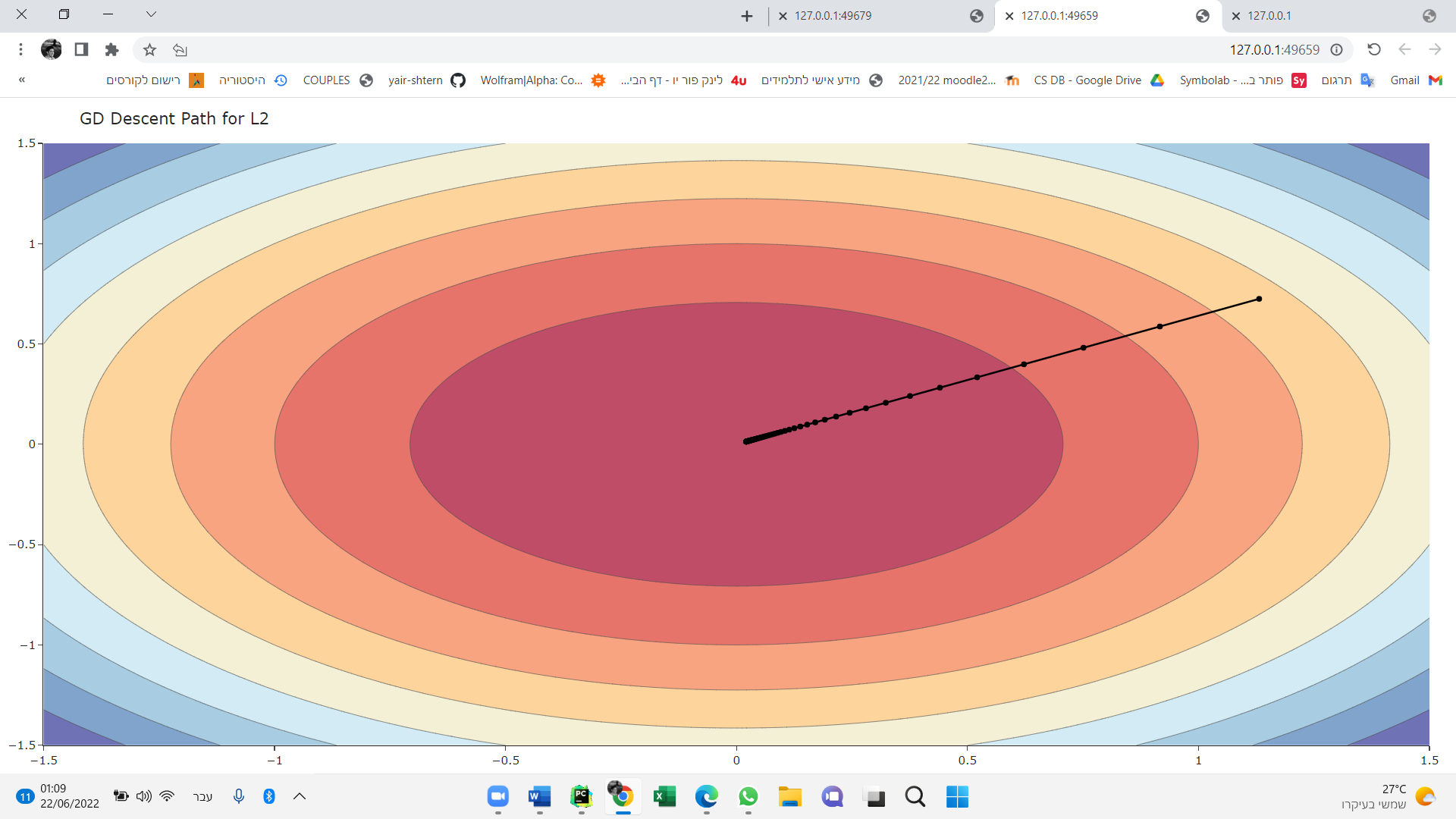
שאלה 7:

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

נשים לב כי שני המקרים בL2 דומים (שניהם מתכנסים למינימום אבל המקרה האקספוננציאלי מתכנס מהר יותר), לעומת המקרה הנוכחי ב־ L1 השונה מהמקרה של קצב הלמידה הקבוע. במקרה של קצב הלמידה הקבוע קיבלנו כי הקצב הקבוע הוביל להתבדרות סביב המינימום, לעומת המקרה הנוכחי בו ניתן לראות שיש התכנסות אל המינימום שכן האלגוריתם מחשב את הצעד הכדאי ביותר בכל איטרציה ובכך ”דילוג“ על 14 הנקודה בה הגרדיאנט משנה את כיוונו (0 = y (לא משנה את ההתכנסות אל המינימום.





שאלה 8:

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי



שאלה 9:

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

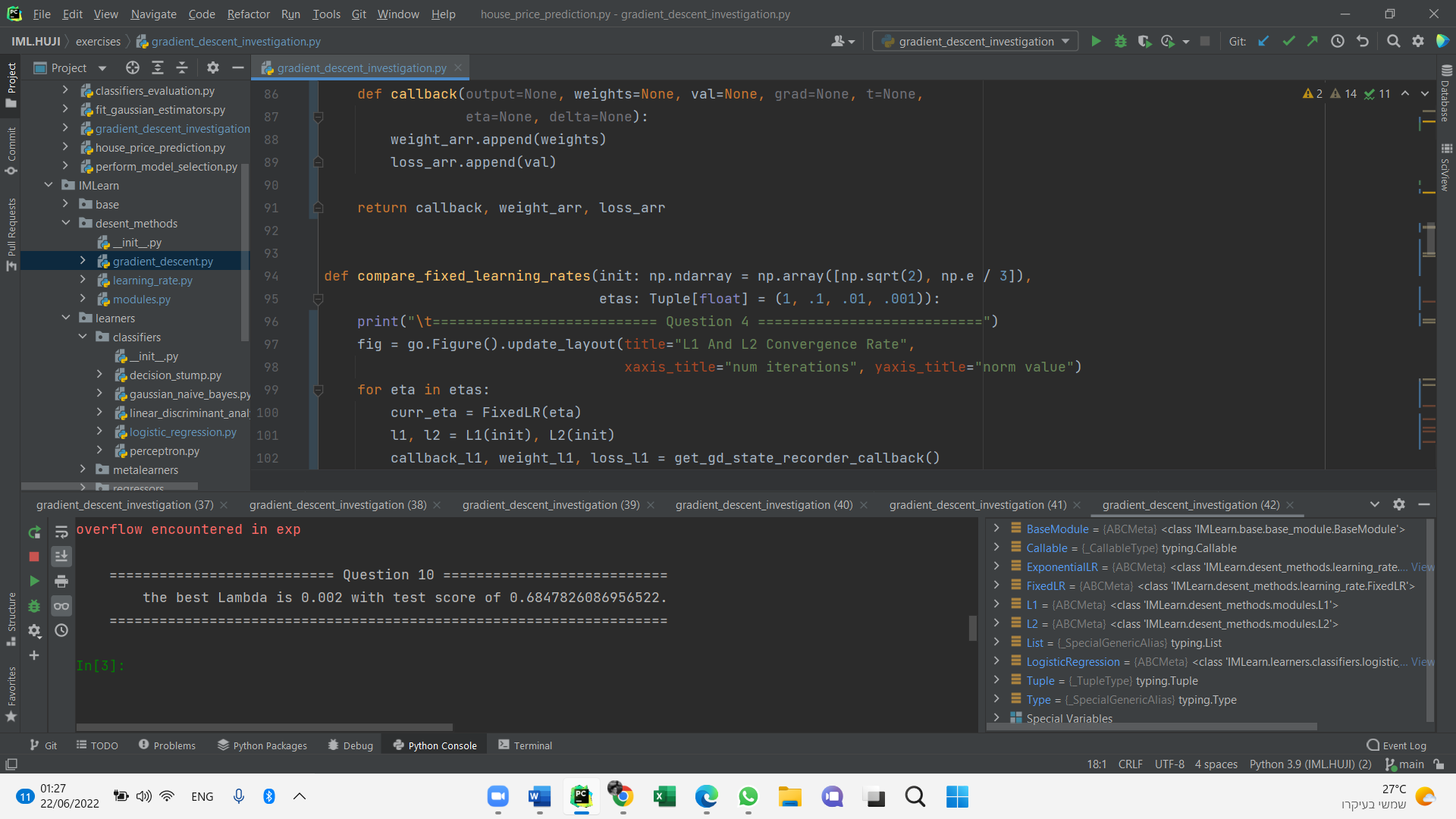
שאלה 10:

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי



שאלה 11:

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי