

תקציר

בעיית החלוקה לאשכולות הינה בעיה בסיסית ומרכזית בתחום הלמידה החישובית הבלתי מונחית, והיא שימושית במגוון רחב של יישומים. בעיית החלוקה לאשכולות היא בעיה לא קמורה ולא חלקה ולפיכך היא קשה לפתירה. בעבודה זו מוצעים שני אלגוריתמים עבור בעיית החלוקה לאשכולות מבוססי מרכזים, כל אלגוריתם פותר את הבעיה עבור פונקצית מרחק שונה. אלגוריתם KPALM פותר את הבעיה כאשר פונקצית המרחק היא המרחק האוקלידי בריבוע, בעוד שאלגוריתם ε -KPALM פותר את הבעיה כאשר פונקצית המרחק היא המרחק האוקלידי הסטנדרטי. שני האלגוריתמים מבוססים על השיטה הקלאסית של מינימיזציה משוחלפת (Alternating Minimization). בשלב חלוקת המדגם, כל אחד משני האלגוריתמים משייך לכל מרכז את הנקודות המתאימות מהמדגם ע"י מציאת מינימום של תת בעיה, שפונקצית המטרה שלה היא ליניארית בתוספת גורם רגולריזציה. בשלב עדכון המרכזים, אלגוריתם KPALM פותר בעיית מינימיזציה בצורה מדויקת, בעוד שאלגוריתם ε -KPALM מוצא פתרון מקורב לבעיית המינימיזציה ע"י ביצוע צעד גרדיאנט. בעבודה מופיעה הוכחת התכנסות גלובלית של האלגוריתמים הנ"ל לנקודה קריטית, הנעשית ע"י שימוש במתודולוגיה חדשנית המבוססת על תכונת Kurdyka-Łojasiewicz. בנוסף לתוצאות התיאורטיות, אנחנו מציגים תוצאות נומריות המדגימות את האפקטיביות של האלגוריתמים המוצעים.

אוניברסיטת תל אביב
הפקולטה למדעים מדויקים ע"ש ריימונד וברלי סאקלר

**מחלקה חדשנית של אלגוריתמים המתכנסים באופן גלובלי
לבעיית האשכול**

תזה זו מהווה חלק מהדרישות לקבלת
התואר "מוסמך למדעי מדויקים" באוניברסיטת תל אביב

אוניברסיטת תל אביב
בית הספר למדעי המתמטיקה
החוג לסטטיסטיקה וחקר ביצועים

מאת סרגיי וולדמן

המחקר נערך בהנחיית
פרופ' מרק טבול
ופרופ' שהם סבח

פברואר 2016