

EXAMENSARBETE Patterns in live performance data**STUDENT** Simon Svensson**HANDLEDARE** Prof. Pierre Nugues (LTH) and Jens Gulin (Sony Moblie)**EXAMINATOR** Jacek Malec (LTH)

Metod för att hitta förklaringar till långsamma app start-tider

POPULÄRVETENSKAPLIG SAMMANFATTNING **Simon Svensson**

Androidmarknaden är en gigantisk marknad som växer för varje dag. Den stora mängden data som genereras måste behandlas på ett bra sätt för att hjälpa utvecklarna att skapa bättre och snabbare program samt telefoner. Här presenteras ett arbetssätt som kan hitta problem med Android-appar och därför hjälpa utvecklare att lösa problem.

Sony Moblie samlar viss prestandadata från användare av telefoner som har tillåtit detta. Denna data innehåller bland annat tiden det tar för en app att starta, tiden från att användare klickar på appen till att den är redo att användas. Denna data har analyserats av prestandaingenjörer med hjälp av klassiska statistiska metoder och andra verktyg. I detta examensarbete har denna data analyseras med hjälp av maskininlärning, en typ av artificiell intelligens som använder stora mängder data för att lära programmet. Detta är för att hitta vad det är som gör starttiden för appar på telefonen långsamma. En metod för att hitta vad som gör telefoner långsamma har implementerats och testas för att använda i framtiden. Detta program kallas "Alarm".

Det första problemet var att representera tidsdata på ett bra sätt. Först testades det att använda tiden som den är till skapandet av modellen. Den modellen fungerade inte bra, det blev väldigt mycket fel i modellen och gav nästintill ingen information. Modellen byggd på denna tidrepresentation kunde inte återskapa datan och lärde sig väldigt lite från datan.

Efter detta testades en metod med att använda ett medelvärde av tider för att representera en hel grupp av data, för att skapa en bredare approxi-

mation av datan.

Detta fungerade mycket bättre och gav ett tydligare resultat. Det enda problemet med att ta medelvärdet är att en stor del av informationen som finns i datan försvinner när endast medelvärdet används.

Därför testade att beskrivas datan med normalfördelningar. Detta gav ett liknade resultat som att använda medelvärdet men, det blev mycket lättare för utvecklare som använder programmet att förstå vad som hade hänt. Detta vidare utvecklades sen till "Alarm" programmet. Det består av fem steg.

1. Förarbete och städning av datan.
2. Hitta en distribution som passar data.
3. Hitta skillnader från tidigare distributioner.
4. Använda linjär regression, en metod för att beskriva datan som en linje, på datan.
5. Analysera linjära regressions resultatet för att hitta vad som har ändrats.

När systemet används görs detta genom att använda två olika dataset och jämföra skillnaderna mellan deras resultat. Via experiment har vi kunnat visa att "Alarm" kan hitta att till exempel en app har blivit mycket långsammare.