

## Populärvetenskaplig sammanfattning

### Anomalidetektion i Strömmande Tidsseriedata med Aktivt Lärande och Metainläring Jonas Lundgren

Mängden strömmande data som finns tillgänglig ökar för varje år som går, mycket på grund av ökad användning av sensorer och Internet of things produkter. När avvikande observationer uppstår i den strömmande datan vill vi upptäcka och uppmärksamma dem. En avvikande observation skulle kunna uppstå för att det är ett fel på en sensor eller för att något håller på att gå snett till. Ett fel i en maskin som är viktig för produktion på ett företag, kan resultera i stora förluster och därför är det viktigt att snarast möjligt upptäcka den här typen av fel. Det finns emellertid många typer av applikationsområden där upptäckter av avvikande datapunkter är viktigt; inom sjukvård, säkerhet och för att upptäcka bedrägeri.

Ett stort problem med att upptäcka avvikande datapunkter är att de är få i antal. Därför finns det inte lika många exempel att träna på. Även om det är ett svårt att upptäcka avvikande observationer finns den maskininlärningsmodeller och algoritmer skapta specifikt för den här typen av uppgift. Dock finns det även problem för den här typen av modeller och algoritmer. Ett sådant är att de först tränas på en uppsättning data för att sedan sättas i produktion. När modellen väl börjat användas så är det svårt att påverka den.

I det här exjobbet har ett ramverk för att upptäcka avvikande observationer skapats med två huvudsakliga mål i åtanke. Det första målet var att ramverket ska vara adaptivt nog att anpassa sig till förändringar i den strömmande datan. Exempelvis kan vi tänka oss att värden från en värmesensor kommer förändras under dag och natt. Det andra målet var att ge ramverket möjlighet att be en människa om hjälp och därefter uppdateras och bli bättre. På så sätt har modellen möjlighet att förändras och förbättras även efter att den är satt i produktion. En expert på området där modellen används har också möjlighet att justera modellen så att den ska prestera bättre.

Resultaten visade att ramverket kan anpassa sig till förändringar i strömmande data samt att det går att påverka ramverkets framtida prestation. Ju mer hjälp ramverket får av en mänsklig expert desto bättre presterar det. Det har testats på ett referensdatasätt och jämförts med hur andra modeller, anpassade för att hitta avvikande observationer, har presterat. Från jämförelserna konstateras att ramverket presterar bättre än vissa och sämre än andra modeller. Dock har inga av modellerna som ramverket jämfördes med möjlighet att förändras baserat på mänsklig påverkan och det är därför inte en helt rättvis jämförelse.