

EXJOBBSFÖRSLAG – 30 HP

Semantisk klassificering av nyhetsartiklar

Bakgrund

Svenska Dagbladet (SvD) är en rikstäckande kvalitetstidning med publicering i både papper och digitala kanaler. SvD ingår i Schibsted Media Group, där även bolag som exempelvis Aftonbladet och Blocket ingår. Det är en internationell mediekoncern med cirka 6 800 anställda i 31 länder.

I konkurrensen om uppmärksamhet från dagens mediekonsumenter är det inte bara publiceringen av relevanta nyheter som avgör framgång för tidningen utan även användarupplevelsen. För att kunna bibehålla vår starka position på marknaden kommer stort fokus framöver ligga på produktutveckling i form av personaliserat innehåll baserat på användarpreferenser och -data. Med fler än 20 miljoner besök per månad på svd.se genereras och bearbetas stora mängder data kopplat till användare, produktens funktionalitet och innehåll. Dessa data utgör grunden för de insikter som genereras.

SvD kommer framöver att fokusera på än mer personaliserat innehåll för att kunna leva upp till konsumenternas nya förväntningar som växer fram i hög takt. Ett led i detta arbete är att bättre förstå vilken typ av artiklar vi dagligen publicerar och vilket användarbeteende artiklarna triggar.

Projektbeskrivning

Ambitionen är att med hjälp av metoder från språkteknologi och text mining ta fram en metod för semantisk klassificering av SvD:s artiklar. Metoden kommer att användas för att skapa rätt förutsättningar för att förstå vilket innehåll som driver olika läsar- och kundbeteenden.

Arbetet kan exempelvis omfatta att ta fram en lämplig semantisk klassificerare för artiklar skrivna på svenska, och att utveckla befintliga topic model-algoritmer genom att inkludera exempelvis ordföljd, ordklass, sentiment och semantik.

Examensarbetet kommer vara en unik chans att få arbeta med ett världsledande digitalt mediehus och hjälpa en av Sveriges största tidningar att ta fram en avancerad datamodell som kommer att kunna tillämpas inom en rad områden.

Vi erbjuder en kreativ och öppen miljö i våra lokaler belägna i centrala Stockholm, en spännande arbetsplats där alla Sveriges Schibstedbolag är samlade under ett och samma tak. Som examensarbetare kommer ni att erbjudas en plats i våra lokaler men önskas arbetet genomföras på distans går även det att lösa, det kommer finnas utrymme för flexibilitet.

Du kommer att arbeta nära SvD:s analysteam som består av tre data scientists och två analytiker som kontinuerligt arbetar med att optimera produkt och affär.

Studentprofil

Vi tror att du som söker har ett intresse av att arbeta med stora datamängder och kännedom om olika text- och data mining-algoritmer. Relevanta kurser är 732A92, TDDE16 och TDDE09. Arbetet underlättas av programmeringsvana i R, SQL samt erfarenhet av AWS eller Spark.

Kontakt och potentiell handledare vid LiU

Marco Kuhlmann, marco.kuhlmann@liu.se