Rekommendations motor:

– Med fokus inom läroverktyg

Examensarbete Systemarkitekturutbildningen

Lennart Jakobsson

Thires Nilsson

HT 2016/VT 2017: NSA012



**Systemarkitekturutbildningen** är en kandidatutbildning med fokus på programutveckling. Utbildningen ger studenterna god bredd inom traditionell program- och systemutveckling, samt en spets mot modern utveckling för webben, mobila enheter och spel. Systemarkitekten blir en tekniskt skicklig och mycket bred programutvecklare. Typiska roller är därför programmerare och lösningsarkitekt. Styrkan hos utbildningen är främst bredden på de mjukvaruprojekt den färdige studenten är förberedd för. Efter examen skall systemarkitekter fungera dels som självständiga programutvecklare och dels som medarbetare i en större utvecklingsgrupp, vilket innebär förtrogenhet med olika arbetssätt inom programutveckling.

I utbildningen läggs stor vikt vid användning av de senaste teknikerna, miljöerna, verktygen och metoderna. Tillsammans med ovanstående teoretiska grund innebär detta att systemarkitekter skall vara anställningsbara som programutvecklare direkt efter examen. Det är lika naturligt för en nyutexaminerad systemarkitekt att arbeta som programutvecklare på ett stort företags IT-avdelning, som en konsultfirma. Systemarkitekten är också lämpad att arbeta inom teknik- och idédrivna verksamheter, vilka till exempel kan vara spelutveckling, webbapplikationer eller mobila tjänster.

Syftet med examensarbetet på systemarkitekturutbildningen är att studenten skall visa förmåga att delta i forsknings- eller utvecklingsarbete och därigenom bidra till kunskapsutvecklingen inom ämnet och avrapportera detta på ett vetenskapligt sätt. Således måste de projekt som utförs ha tillräcklig vetenskaplig och/eller innovativ höjd för att generera ny och generellt intressant kunskap.

Examensarbetet genomförs vanligen i samarbete med en extern uppdragsgivare eller forskningsgrupp. Det huvudsakliga resultatet utgörs av en skriftlig rapport på engelska eller svenska, samt eventuell produkt (t.ex. programvara eller rapport) levererad till extern uppdragsgivare. I examinationen ingår även presentation av arbetet, samt muntlig och skriftlig opposition på ett annat examensarbete vid ett examinationsseminarium. Examensarbetet bedöms och betygssätts baserat på delarna ovan, specifikt tas även hänsyn till kvaliteten på eventuell framtagen mjukvara. Examinator rådfrågar handledare och eventuell extern kontaktperson vid betygssättning.



Besöksadress: Järnvägsgatan 5 · Postadress: Allégatan 1, 501 90 Borås

Tfn: 033-435 40 00 · E-post: inst.hit@hb.se · Webb: www.hb.se/hit

**Svensk titel:** Rekommendations motor: med fokus inom läroverktyg

**Engelsk titel:** Recommendation engine: focus within tools for learning

**Utgivningsår:** 2017

**Författare:** Lennart Jakobsson & Thires Nilsson

**Handledare:** Tuwe Löfström

**Abstract**

(på engelska)

**Keywords:** (på engelska)

**Sammanfattning**

(på svenska)

**Nyckelord:** (på svenska)

Innehållsförteckning

[1 Inledning - 1 -](#_Toc468371862)

[1.1 Någon passande rubrik - 1 -](#_Toc468371863)

[1.2 Problemdiskussion: Verksamhetsanpassa en rekommendationsmotor - 2 -](#_Toc468371864)

[1.3 Problemfomulering & Rekommendationsmotorns syfte - 4 -](#_Toc468371865)

[1.4 Bakgrund - 4 -](#_Toc468371866)

[1.5 Syfte - 4 -](#_Toc468371867)

[2 Referenser - 5 -](#_Toc468371868)

# Inledning

Information som finns tillgänglig på internet växer ständigt, utifrån detta blir det svårare och svårare för användare att snabbt hitta relevant information. Som resultat växer även behovet för intelligenta och pålitliga rekommendations motorer, inte enbart för merförsäljning utan även i form av lärande medel. Läroverktyg för undervisning via internet har under senare tid funnit sin tjusning och flertalet verksamheter har börjat att erbjuda distanskurser för olika typer av målgrupper. Nomp är ett av dessa företag vars verksamhet bygger på att erbjuda ett smidigt läroverktyg för barn och ungdomar som hjälpmedel för inlärning av matematik. Tidigare studier bygger bland annat på hur man kan ge pålitliga prediktioner och hur olika typer av rekommendationssystem fungerar. Eftersom fokus för denna typ av verksamhet inte direkt ligger på merförsäljning utan snarar användarens lärande leder detta till ökat behov av rekommendationsmotorer som tar detta i beaktning. Målsättningen med studien är därför att undersöka skillnader mellan rekommendationsmotorer i utbildningssyfte respektive merförsäljningssyfte samt tekniker för att implementera en rekommendations motor anpassad för verksamhet med fokus inom lärande.

## Bakgrund

Rekommendationssystem finns i dagsläget i stort utbud på många olika typer av webbplatser. Även om de har liknande syfte så kan implementeringen av dessa skilja mycket då det finns flera typer av algoritmer för att lösa detta problem. Val av algoritm baseras och kan även anpassas utefter verksamhetens produkter och/eller tjänster, vilket ofta resulterar i dessa skillnader då ingen verksamhet är den andre lik.

En del av de algoritmer som studerats mycket inom detta område redogörs av Gorakala och Usuelli (2015) vilka är: ”Collaborative filtering”, ”Content-based recommending ”, ”Knowledge-based recommending” samt ”Hybrid approaches”.

Collaborative i engelskans betydelse är samarbete vilket också var grundidén till denna algoritm. Att genom samarbete mellan användare kunna rekommendera produkter mellan andra användare som delar samma intresse. Gorakala och Usuelli (2015) förklarar vidare att om viss person A och person B gillar exempelvis samma bok så finns det en stor sannolikhet att dessa delar samma intresse i framtiden. Om person A då skulle köpa en ny bok så skulle därför en rekommendation på denna bok kunna göras för person B.

”Content-based” rekommendations system skiljer sig genom att ta likheter mellan produkter i beaktning. För att åstadkomma detta behövs en relation mellan vissa produkter exempelvis produkters olika kännetecken, i bokexemplet kan detta förslagsvis vara bokens genre. Utifrån tidigare visat intresse av en användare för en specifik genre kan man därför rekommendera andra böcker av liknande genre. Prediktioner görs alltså inte genom jämförelser mellan andra användare utan enbart den enskilde användarens tidigare preferenser (Aggarwal, 2016)

”Kowledge-based” rekommendations system används enligt Gorakala och Usuelli (2015) i mer specifika domäner där användarens historiska data är något mer begränsad. Detta kan ske i verksamheter där många användare är helt nya användare och det därför inte finns någon tidigare data, alltså ingen information angående preferenser. Metoden löser alltså problemet med bristfällig information om preferenser genom att inhämta ytterligare information från användaren. Liknande ”Content-based” rekommendations system krävs även här kännetecknen(features) för olika produkter vilket jämförs med användarens preferens och därefter rekommenderas.

Utöver dessa algoritmer så finns hybrida metoder vilket resulterar i ett mer robust system (Gorakala & Usuelli, 2015).Tanken med detta är att kombinera olika egenskaper av de tidigare förklarade metoderna. Svagheter hos en metod kan därför förstärkas eller helt elimineras genom att kombinera funktionalitet hos en annan metod. En svaghet med tillexempel ”Collaborative-filtering” är då nya produkter läggs till i systemet och då historik för denna produkt saknas. Vilket leder till att produkten sällan rekommenderas eller inte alls. Genom att då kombinera funktionalitet från ”Content-based” metoden kan man rekommendera en helt ny produkt till användare som köpt eller intresserat sig för produkter med liknande kännetecknen (features).

Sammanfattningsvis är verksamheten avgörande för vilken metod/algoritm som bör användas som implementering för rekommendationsmotorn. Vad som lämpligast passar en verksamhet som säljer filmer via internet kanske inte passar en applikation som rekommenderar restauranger för potentiella kunder. Anledningen till detta är att i båda fall behöver man vara tydlig med vad som är till intresse för användare. En rekommendationsmotor som rekommenderar restaurangen behöver exempelvis ta avstånd i beaktning då en potentiell kund i majoritetsfall inte färdas orimligt långt för en måltid enligt preferenser (Aggarwal, 2016).

## Problemdiskussion: Verksamhetsanpassa en rekommendationsmotor

För att kunna anpassa en rekommendationsmotor till en verksamhet så behöver man framförallt fastställa vad det verkliga syftet med rekommendationerna är. I många fall där verksamheter sysselsätter sig med olika typer av försäljning är syftet troligtvis ganska ofta merförsäljning. Den verksamhet som ligger till grund för denna studie är en verksamhet som fokuserar på barn och ungdomars lärande i matematik. Givetvis är inte merförsäljning den huvudsakliga fokus som driver denna typ av verksamhet och samma lösning kan därför inte användas som för de mer typiska rekommendationsmotorerna.

I verksamheten ingår både elever, lärare och även föräldrar som stöd för elevernas lärande. Lärare kan tilldela elever uppgifter och föräldrar kan ha en bättre översikt av barnets framgång. Lärarens intresse är självklart att se att dess elever klarar av de uppgifter som delas ut. Dock har även lärare ibland riktlinjer att ta hänsyn till, exempelvis läroplaner och dylikt. Elevernas intresse är självklart att klara de uppgifter som blivit tilldelade av lärare och möjligtvis att lära sig mer alternativt öva tidigare erhållen kunskap. Då intresset hos eleverna skiljer sig betydligt från lärarna behöver rekommendationsmotorn antingen ta detta i beaktning eller bara hantera en specifik användargrupp.

På grund av lärares olika förhållningssätt att lära vad gäller metodik och/eller planering så bör därför en rekommendationsmotor ha mer syfte och betydelse förslagsvis hos eleverna. Därför avgränsas rekommendationsmotrons hantering till endast eleverna.

Eleverna kan som tidigare nämnt få tilldelade uppgifter av lärare men det finns även barn och ungdomar vars lärare inte använder detta verktyg. Dessa elever saknar riktlinjer så som tilldelade uppgifter och för denna målgrupp hade en rekommendationsmotor passat väldigt bra. Men självklart blir även rekommendationer från systemet ett extra verktyg för alla elever att lära sig mer.

Förutom att ta hänsyn till användargrupp behöver man även fastställa vad som för i detta fall eleverna utgör en bra rekommendationsmotor. Utifrån Aggarwal’s (2015) förklaring på vad som utgör en riktigt bra rekommendationsmotor så behöver den uppfylla vissa egenskaper vilka han namnger enligt följande: Relevans, nymodighet (”Novelty”), Lyckoträff (”Serendipity”), variation i rekommendationerna

Relevans är en av de mer självklara egenskaperna för en rekommendationsmotor, att givetvis rekommendera sådant som för användaren är relevant. I elevernas fall behöver vi för att kunna lösa detta besvara vad som faktiskt är relevant. Gör eleven ett par uppgifter inom samma matematiska kategori och gör dessa utan några som helst svårigheter kan det vara relevant att få rekommenderat en något svårare uppgift. Å andra sidan om eleven har svårigheter med dessa uppgifter kan det vara relevant att rekommendera en något lättare uppgift. Det kan även finnas relevans i att rekommendera uppgifter för en elev A som inte har en registrerad lärare men har gjort ett antal uppgifter som även gjorts av elev B. Med skillnaden att elev B har en registrerad lärare och en tilldelad uppgift som innehåller dessa uppgifter, och därför rekommendera resterande tilldelade uppgifter som elev B har för elev A.

Tanken med nymodighet som Aggarwal (2015) beskriver det är att föreslå produkter och/eller tjänster för användaren som den inte sett eller varit medveten om tidigare. I denna typ av system så är det inte särskilt meningsfullt att rekommendera uppgifter för eleven som den tidigare gjort utan bör därför till bästa förmåga endast rekommendera sådana uppgifter som eleven ännu inte slutfört. Däremot kan elever ha påbörjat uppgifter men inte slutfört dessa, rekommendationsmotorn bör därför kunna rekommendera uppgifter som eleverna har påbörjat men inte slutfört. Anledningen till detta kan exempelvis vara att eleven tidigare inte besatt den kunskapen som krävdes för att lösa uppgiften. Eleven kan ha gjort andra uppgifter för att fördjupa sina kunskaper eller dylikt. Om eleven löser liknande uppgifter relativt enkelt bör icke slutförda uppgifter därför kunna rekommenderas av systemet.

”Serendipity” eller lyckoträff som det kan översättas till på svenska, handlar om att rekommendera sådant som för användaren kommer lite som en överraskning. I denna typ av system kan det vara svårare att åstadkomma då domänen förhållandevis är ganska begränsad. I en annan domän exempelvis om man sysslar med försäljning av böcker kan man även kolla på vad användaren har tittat på och inte nödvändigtvis läst. Baserat på detta göra rekommendationer, vilket då kan komma som en överraskning för användaren och kanske eventuellt locka användaren till att köpa boken. I Nomps fall kan möjligtvis en viss ”Serendipity” också uppnås genom att rekommendera uppgifter för icke vägledda elever. Eftersom att detta kan uppfattas som en form av vägledning så kan möjligtvis känslan av en ”lyckoträff” uppstå. Men att uppnå en lika överraskande känsla som i bokexemplet för alla typer av användare är förmodligen mycket svårare och man behöver då noggrannare undersöka hur man skulle kunna genomföra detta. Att basera rekommendationer tillexempel baserat på vad användaren tidigare har kollat på kan helt sakna relevans om uppgifterna är på tok för svåra för eleven.

Till sist nämner Aggarwal (2015) att det bör finnas skillnad i rekommendationerna. Detta på grund av att om man ger ett antal liknande rekommendationer för användaren så finns sannolikheten att eleven kan finna nytta antingen av alla rekommendationer, eller ingen alls. Har man en viss skillnad på de förslag som ges för användaren så ökar sannolikheten att åtminstone en av uppgifterna är till nytta för eleven. Detta skulle i rekommendationsmotorn åstadkommas genom att i de fall data finns tillgänglig rekommendera uppgifter från olika typer av matematiska områden, exempelvis rekommendera en viss uppgift inom algebran och en viss uppgift inom geometrin.

Syftet med att fastställa vad som krävs av en bra rekommendationsmotor är för att det inte är helt trivialt att på ett säkert sätt säga att de prediktioner som görs för användaren faktiskt är bra. En rekommendationsmotor som rekommenderar produkter som helt saknar relevans för användaren har ingen som helst betydelse. Det finns även andra faktorer som kan vara avgörande för pålitliga prediktioner, faktorer som exempelvis kan vara tid. Om en användare inte har varit aktiv inom systemet under en längre period så kan prediktionerna för användaren vara bra utifrån historiken, men på grund av den tid som gått inte längre vara relevant för eleven. Om det har gått väldigt lång tid kan eleven möjligtvis vara i flera klasser högre än senast.

Utmaningar ligger också i att identifiera den data som är av intresse för systemet att basera sina rekommendationer på, även mängden data kan vara ett problem. Vissa åtgärder kan behöva vidtas för att begränsa den data som systemet använder för att göra prediktioner, beroende på exempelvis exekveringstid och relevans.

## Problemfomulering och frågeställning

Utifrån tidigare redogörelse kring olika typer av metoder/algoritmer samt vad som utgör en bra rekommendationsmotor så framkommer olika typer av frågeställningar som kommer ligga till underlag för att besvara studiens slutliga frågeställning.

Dessa frågeställningar är följande:

Vilken algoritm/metod bör med störst fördel väljas och som även passar problemet/verksamheten?

Vilken typ av data behöver systemet ta hänsyn till för att göra passande prediktioner till systemets användare?

Hur kan vi säkerställa att de prediktioner som görs av systemet är till nytta för användarna?

Skiljer sig kvalitén i prediktionerna genom den metod som valts i jämförelse med en mindre anpassad algoritm?

Studien avgränsas till att anpassa rekommendationsmotorn utifrån nytta för eleverna, rekommendationer för lärare lämnas utanför studien.

Det övergripande syftet med studien är alltså att undersöka möjligheterna till att anpassa en rekommendationsmotor för en verksamhet vars huvudsakliga fokus är lärande via internet, verksamheten i detta fall är Nomp. Syftet är även att implementera rekommendationsmotorn för att genomföra experiment som stöd till studien. Experimenten kommer att vara till underlag för att besvara de delfrågor som när besvarade kommer att underlätta att besvara den slutgiltiga frågeställningen, vilket är:

Hur kan man anpassa samt implementera en rekommendationsmotor för en verksamhet vars huvudsakliga fokus är lärande och som också har nytta för användarna samt verksamheten?

# Referenser

Aggarwal, C. C. (2016). *Recommender systems.* New York: Springer International Publishing.

Gorakala, S. K., & Usuelli, M. (2015). *Building a recommendation system with R.* Birmingham: Packt Publishing Ltd.

**Högskolan i Borås** är en modern högskola mitt i city. Vi bedriver utbildningar inom ekonomi och informatik, biblioteks- och informationsvetenskap, mode och textil, beteendevet­enskap och lärarutbildning, teknik samt vårdvetenskap.

På **institutionen Handels- och IT-högskolan (HIT)** har vi tagit fasta på studenternas framtida behov. Därför har vi skapat utbildningar där anställningsbarhet är ett nyckelord. Ämnesintegration, helhet och sammanhang är andra viktiga begrepp. På institutionen råder en närhet, såväl mellan studenter och lärare som mellan företag och utbildning.

Våra **ekonomiutbildningar** ger studenterna möjlighet att lära sig mer om olika företag och förvaltningar och hur styrning och organisering av dessa verksamheter sker. De får även lära sig om samhällsutveckling och om organisationers anpassning till omvärlden. De får möjlighet att förbättra sin förmåga att analysera, utveckla och styra verksamheter, oavsett om de vill ägna sig åt revision, administration eller marknadsföring. Bland våra **IT-utbildningar** finns alltid något för dem som vill designa framtidens IT-baserade kommunikationslösningar, som vill analysera behov av och krav på organisationers information för att designa deras innehållsstrukturer, bedriva integrerad IT- och affärsutveckling, utveckla sin förmåga att analysera och designa verksamheter eller inrikta sig mot programmering och utveckling för god IT-användning i företag och organisationer.

**Forskningsverksamheten** vid institutionen är såväl professions- som design- och utvecklingsinriktad. Den övergripande forskningsprofilen för institutionen är handels- och tjänsteutveckling i vilken kunskaper och kompetenser inom såväl informatik som företagsekonomi utgör viktiga grundstenar. Forskningen är välrenommerad och fokuserar på inriktningarna affärsdesign och Co-design. Forskningen är också professionsorienterad, vilket bland annat tar sig uttryck i att forskningen i många fall bedrivs på aktionsforskningsbaserade grunder med företag och offentliga organisationer på lokal, nationell och internationell arena. Forskningens design och professionsinriktning manifesteras också i InnovationLab, som är institutionens och Högskolans enhet för forskningsstödjande systemutveckling.



Besöksadress: Järnvägsgatan 5 · Postadress: Allégatan 1, 501 90 Borås

Tfn: 033-435 40 00 · E-post: inst.hit@hb.se · Webb: www.hb.se/hit