

# REKOMMENDATIONS MOTOR: – MED FOKUS INOM LÄROVERKTYG

Examensarbete Systemarkitekturutbildningen

Lennart Jakobsson  
Thires Nilsson

HT 2016/VT 2017: NSA012



UNIVERSITY OF BORÅS  
SCHOOL OF BUSINESS AND IT

**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 förtroende 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 framtagna mjukvara. Examinator rådfrågar handledare och eventuell extern kontaktperson vid betygssättning.



**HÖGSKOLAN I BORÅS**  
INSTITUTIONEN HANDELS- OCH IT-HÖGSKOLAN

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 -
1.1	Bakgrund.....	- 1 -
1.2	Problemdiskussion: Verksamhetsanpassa en rekommendationsmotor .....	- 3 -
1.3	Problemformulering och frågeställning.....	- 4 -
2	Referenser.....	- 6 -

# 1 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 rekommendationsmotorer, inte enbart för merförsäljning utan även i form av läroverktyg för undervisning via internet. Dessa har under senare tid funnit sin tjugning 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 inläring av matematik. Eftersom fokus för denna typ av verksamhet inte direkt ligger på merförsäljning utan snarare användarens lärande leder detta till ökat behov av rekommendationsmotorer som tar detta i beaktning.

Tidigare studier bygger bland annat på hur man kan ge pålitliga prediktioner och hur olika typer av rekommendationssystem fungerar (Herlocker, Konstan, Borchers, & Riedl, 1999; Adomavicius & Tuzhilin, 2005). Målsättningen med studien är därför att undersöka skillnader mellan rekommendationsmotorer i utbildningssyfte respektive merförsäljningssyfte samt tekniker för att implementera en rekommendationsmotor anpassad för verksamhet med fokus inom lärande.

## 1.1 Bakgrund

I Nomp's verksamhet 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 (Skolverket, u.å). Elevernas intresse är självklart att klara de uppgifter som blivit tilldelade av lärare och möjligtvis att lära sig mer alternativt öva tidigare erhållen kunskap. Det finns dock barn och ungdomar vars lärare inte använder detta verktyg, som av den anledningen saknar samma typ av vägledning.

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 andra lik.

En del av de algoritmer som studerats mycket inom detta område redogörs av Gorakala och Usuelli (2015) vilka är: *samarbetsfiltrering* (Collaborative filtering), *innehållsbaserad rekommendering* (Content based filtering), *kunskapsbaserad filtrering* (Knowledge based filtering) samt *hybrida metoder* (Hybrid approaches).

Samarbetsfiltrering var som namnet låter 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 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.

Innehållsbaserade rekommendationssystem 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)

Kunskapsbaserade rekommendationssystem 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 innehållsbaserade rekommendationssystem 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 minimeras eller helt elimineras genom att kombinera funktionalitet hos en annan metod. En svaghet med till exempel samarbetsfiltrering är då nya produkter läggs till i systemet och då historik för denna produkt saknas. Det leder till att produkten sällan eller inte alls rekommenderas. Genom att då kombinera funktionalitet från innehållsbaserad filtrering kan man rekommendera en helt ny produkt till användare som köpt eller intresserat sig för produkter med liknande kännetecken (features).

Bortsett från metod behöver man även fastställa vad som för användarna 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) samt *variation i rekommendationerna*.

Relevans är en av de mer självklara egenskaperna för en rekommendationsmotor, att givetvis rekommendera sådant som för användaren är relevant. Nymodighet enligt Aggarwal's (2015) beskrivning är att föreslå produkter och/eller tjänster för användaren som den inte sett eller varit medveten om tidigare. Lyckoträff handlar om att rekommendera sådant som för användaren kommer lite som en överraskning. Till sist nämner Aggarwal (2015) även 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 rekommendationerna kan vara till nytta.

Sammanfattningsvis 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 exempelvis att en rekommendationsmotor som rekommenderar restaurangen behöver ta avstånd i beaktning. En potentiell kund i majoritetsfall färdas inte orimligt långt för en måltid enligt preferenser (Aggarwal, 2016). 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 nödvändigtvis det huvudsakliga fokus som driver denna typ av verksamhet och samma lösning kan därför inte användas som för de mer typiska rekommendationsmotorerna. Verksamheten är alltså avgörande för vilken metod/algoritm som bör användas som implementering för rekommendationsmotorn.

## 1.2 Problemdiskussion: Verksamhetsanpassa en rekommendationsmotor

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, men kanske framförallt läroplanerna (Skolvereket, u.å) så bör därför en rekommendationsmotor ha mer syfte och betydelse förslagsvis hos eleverna. Att anpassa rekommendationsmotorn utefter lärarnas läroplaner kan möjligtvis leda till ett bekvämt verktyg för lärarna, men enligt vår uppfattning bör rekommendationsmotorn uppfylla mer nytta om den anpassas för eleverna. Den användargrupp som särskilt hade haft stor nytta av en rekommendationsmotor är de elever som saknar vägledning av lärare. Men troligtvis blir även rekommendationer från systemet ett extra verktyg för alla elever att lära sig mer. Av dessa anledningar avgränsas rekommendationsmotorns hantering till endast eleverna.

Utifrån tidigare förklaring vad som utgör en bra rekommendationsmotor så bör vi ha som målsättning att det rekommendationsmotorn rekommenderar har relevans, nymodighet, lyckoträff samt variation. Vilket i sig kan vara en utmaning. För att uppnå relevans kan man möjligtvis utgå ifrån elevens tidigare slutförda uppgifter. Har dessa slutförts 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 att rekommendera uppgifter för elever baserad på tidigare utdelade uppgifter från lärare, speciellt för ej vägleda elever. Det är inte heller 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. Om eleven löser liknande uppgifter relativt enkelt bör icke slutförda uppgifter därför kunna rekommenderas av systemet.

Både lyckoträff och nymodighet i denna typ av system kan vara svårare att åstadkomma då domänen är ganska begränsad. I en annan domän, som exempelvis sysslar med försäljning av produkter kan man ibland ha hundratusentals olika produkter. Vilket gör att sannolikheten för att användaren sett eller hört talas om produkten kan vara relativt liten, och därför lättare överraska användaren. I Nomps fall har vi ett begränsat antal uppgifter och en naturlig ordning till uppgifterna, vilket gör det svårare att överraska en användare.

Variation i rekommendationerna skulle kunna åstadkommas genom att i de fall data finns tillgänglig rekommendera uppgifter från olika typer av matematiska områden, exempelvis rekommendera en viss uppgift i algebran och en viss uppgift i geometrin inom elevens aktuella årskurs.

Den största utmaningen kommer ligga i att på något sätt säga att de prediktioner som görs av systemet faktiskt är bra, detta är inte helt trivialt. En rekommendationsmotor som rekommenderar produkter som helt saknar relevans för användaren har givetvis 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.



Problemet 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.

### 1.3 Syfte och frågeställning

Det övergripande syftet med studien är 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 den slutgiltiga frågeställningen, vilket är:

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

Utifrån tidigare redogörelse kring olika typer av metoder/algoritmer samt vad som utgör en bra rekommendationsmotor så framkommer olika delfrågor vars svar kan underlätta för att besvara studiens slutliga frågeställning.

Dessa delfrågor är följande:

Vilken algoritm/metod bör väljas som passar problemet/verksamheten?

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

Skiljer sig kvalitén i prediktionerna genom den metod som valts i jämförelse med ett redan befintligt alternativ?

## 2 Metod

Data som är nödvändig för utförandet av denna studie är bland annat historisk från användarna i Nomp's system. Variabler som kan vara av intresse är exempelvis uppgifter respektive användare gjort, antalet fel eleven har gjort i slutförandet av uppgiften samt avlagd tid. Studien kommer även att innefatta implementering av rekommendationssystemet. Utifrån vald implementering kommer resultat från testning av systemet motsvarar empirin för denna studie vilket också kommer att utgöra grund för att påvisa nyttan med systemet.

### 2.1 Forskningsdesign

Vi har valt att arbeta enligt metoden *DSR* (Design science research). Denna metod passar studien då vi, för att besvara forskningsfrågan bäst bör kunna göra detta med en anpassad implementation av ett rekommendationssystem. Vi kommer att utgå ifrån den modell som utvecklades av Takeda, Veerkamp, Tomiyama, och Yoshikawa (1990) som även kallas *DSR* cykeln (Vaishnavi & Kuechler, 2015). Enligt metodens fasar bygger den fjärde på att man ska evaluera artefakten man implementerat för påvisa nyttan med systemet. För att evaluera rekommendationssystemet kommer vi att använda oss av kända metoder för att mäta träffsäkerheten i prediktionerna. Exempel på några välkända metoder för att evaluera träffsäkerheten är *Spearman rank correlation coefficient* eller *Kendall rank correlation coefficient* (Aggarwal, 2016).

De rekommendationer som systemet kommer basera prediktionerna på utgörs av dataset framställda från Nomp's användares historik. Samma dataset kommer att användas för att mäta träffsäkerheten i prediktioner för rekommendationsmotorn vi implementerat samt jämföras med en mindre anpassad lösning alternativt en redan välkänd lösning. Syftet att jämföra träffsäkerheten i prediktionerna med en annan redan känd lösning är att få ett

mätvärde som tydligt påvisar nyttan med det implementerade systemet. Detta experiment kommer att upprepas på andra dataset för att få ett mer generellt värde, samt utesluta möjligheten att en viss implementation presterar bättre på ett specifikt dataset. Jämförelser med den verkliga datan exempelvis uppgifter som lärare typiskt brukar gruppera tillsammans kan även ge en hänvisning om rekommendationerna har relevans.

Eftersom att mätvärdena blir förhållandevis tydligt en variabel lämpar sig denna studien också mot en kvantitativ ansats men kommer utgå ifrån DSR-ramverket.

För att få mer underlag för nyttan i systemet kommer intervju med en kvalificerad mattelärare bekant med systemet genomföras. Matteläraren kommer att få exempel på prediktioner som systemet gjort och granska dessa. Baserat på resultat från intervju samt empirin kommer analys av detta ske med utgångspunkt i vad som utgör en bra rekommendationsmotor.

Då studien kräver att vi implementerar artefakten för att kunna utföra de experiment studien bygger på så finns det inte många alternativa metoder som hade passat studien. Det alternativ som hade kunnat vara relevant är *Action Research*. Men enligt Järvinen (2007) är båda dessa metoder väldigt lika. Vidare hävdar han att det som bland annat verkar skilja metoderna åt är i hur metoderna förhåller sig till verkligheten. Där DSR utgår ifrån verkligheten och försöker lösa problem som existerar i den medan i Action Research modifierar man verkligheten eller utvecklar ett nytt system. Då vi utgår ifrån existerande forskning inom rekommendationssystem men vill bidra med vidare kunskap inom andra typer av domäner så tror vi DSR passar studien bättre.

## 2.2 Metodreflektion

Metodvalet passar studien väl, då studien kräver att vi implementerar en tänkt lösning för att utföra våra experiment så är det egentligen ett självklart val. Då vi får ett ganska tydligt mätvärde utifrån pricksäkerheten i prediktionerna finns så blir dessa förmodligen lätta att analysera. Nackdelen med metoden är att det verkar finnas många olika förhållningssätt inom DSR. Andra metoder inom exempelvis kvantitativ och kvalitativ forskning är mer utstakade vilket kan kännas mindre förvirrande. Som tidigare nämnt planerar vi att öka tillförlitligheten i studien genom bland annat intervju med en kvalificerad lärare inom matematik.

### 3 Referenser

- Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. (IEEE, Red.) *IEEE Transactions on Knowledge and Data Engineering*, 17(6), ss. 734-749.
- Aggarwal, C. C. (2016). *Recommender systems*. New York: Springer International Publishing.
- Gorakala, S. K., & Usielli, M. (2015). *Building a recommendation system with R*. Birmingham: Packt Publishing Ltd.
- Herlocker, J. L., Konstan, J. A., Borchers, A., & Riedl, J. (1999). An algorithmic framework for performing collaborative filtering. *SIGIR '99 Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval* (ss. 230-237). Berkeley: ACM New York. doi:10.1145/312624.312682
- Skolvereket. (u.å). *Matematik*. Hämtat från Skolverkets hemsida: <http://www.skolverket.se/laroplaner-amnen-och-kurser/grundskoleutbildning/grundskola/matematik#anchor1> den 08 12 2016
- Takeda, H., Veerkamp, P., Tomiyama, T., & Yoshikawa, H. (1990). Modeling Design Processes. *AI Magazine*, 11, ss. 37-48.
- Vaishnavi, V. K., & Kuechler, W. J. (2015). *Design Science Research Methods and Patterns, innovating information and communication technology* (2nd edition uppl.). Boca Raton: CRC Press.

**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 forskningsstödande systemutveckling.



**HÖGSKOLAN I BORÅS**

VETENSKAP FÖR PROFESSION

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