

## Text Mining

Text Mining  
6 hp

Programkurs

TDDE16

Gäller från: 2025 VT

<b>Fastställd av</b>	<b>Huvudområde</b>
Programnämnden för data- och medieteknik, DM	Informationsteknologi, Datateknik, Datavetenskap
<b>Fastställandedatum</b>	<b>Utbildningsnivå</b> <b>Fördjupningsnivå</b>
2024-08-28	Avancerad nivå      A1F
<b>Reviderad av</b>	<b>Utbildningsområde</b>
	Tekniska området
<b>Revideringsdatum</b>	<b>Ämnesgrupp</b>
	Datateknik
<b>Gavs första gången</b>	<b>Gavs sista gången</b>
VT 2017	
<b>Institution</b>	<b>Ersätts av</b>
Institutionen för datavetenskap	

## Kursen ges för

- Civilingenjörsprogram i datateknik
- Civilingenjörsprogram i informationsteknologi
- Civilingenjörsprogram i mjukvaruteknik
- Civilingenjörsprogram i teknisk fysik och elektroteknik
- Civilingenjörsprogram i teknisk fysik och elektroteknik - internationell
- Masterprogram i datavetenskap

## Rekommenderade förkunskaper

- diskret matematik, analys, linjär algebra, sannolikhetslära
- programmering, datastrukturer och algoritmer
- maskininlärning

Kursens laborationsserie använder programspråket Python.

## Lärandemål

Text mining utvecklar metoder för att komma åt information i och extrahera kunskap från stora mängder text. Denna kurs har som mål att ge praktisk erfarenhet av de huvudsakliga stegen i text mining: informationsökning, bearbetning av textdata, modellering, analys av experimentella resultat. Kursen avslutas med ett individuellt projektarbete där de studerande arbetar med en egendefinierad frågeställning.

Efter genomgången kurs ska den studerande kunna:

1. implementera text mining-metoder och tillämpa dem på praktiska problem
2. analysera och sammanfatta resultat från text mining-experiment
3. identifiera, formulera och lösa problem inom text mining-området
4. klart redogöra för och diskutera slutsatserna av ett projektarbete

## Kursinnehåll

- informationsökning
- grundläggande språkteknologiska metoder
- prediktiv modellering, särskilt textklassificering
- textklustring och temamodellering
- informationsextraktion
- valideringsmetoder
- tillämpningar av text mining
- programbibliotek och data

## Undervisnings- och arbetsformer

Kursen ges i form av föreläsningar, datorlaborationer och handledning i samband med ett individuellt projektarbete.

## Examination

PRA1	Projekt	3 hp	U, 3, 4, 5
LAB1	Datorlaborationer	3 hp	U, G

Kursens slutbetyg är betyget på PRA1.

## Betygsskala

Fyrgradig skala, LiU, U, 3, 4, 5

## Övrig information

### Påbyggnadskurser

Språkteknologi

### Om undervisnings- och examinationsspråk

Undervisningsspråk visas på respektive kurstillfälle på fliken "Översikt".  
Examinationsspråk relaterar till undervisningsspråk enligt nedan:

- Om undervisningsspråk är "Svenska" kan kursen ges i sin helhet på svenska eller delvis på engelska. Examinationsspråk är svenska, men delar av examinationen kan ske på engelska.
- Om undervisningsspråk är Engelska ges kursen i sin helhet på engelska. Examinationsspråk är engelska.
- Om undervisningsspråk är "Svenska/Engelska" ges kursen i sin helhet på engelska om studenter utan tidigare kunskap i svenska språket deltar. Examinationsspråk följer undervisningsspråk.

### Övrigt

Kursen bedrivs på ett sådant sätt att likvärdiga villkor råder med avseende på kön, könsöverskridande identitet eller uttryck, etnisk tillhörighet, religion eller annan trosuppfattning, funktionsnedsättning, sexuell läggning och ålder.

Planering och genomförande av kurs skall utgå från kursplanens formuleringar. Den kursvärdering som ingår i kursen skall därför genomföras med kursplanen som utgångspunkt.

Kursen är campusförlagd på den ort som anges för kurstillfället om inget annat anges under "Undervisnings – och arbetsformer". I en campusförlagd kurs kan dock enstaka moment på distans ingå.

## Generella bestämmelser

### Kursplan

För varje kurs ska en kursplan finnas. I kursplanen anges kursens mål och innehåll samt de särskilda förkunskaper som erfordras för att den studerande skall kunna tillgodogöra sig undervisningen.

### Schemaläggning

Schemaläggning av programkurser görs enligt, för kursen, beslutad blockindelning. Fristående kurser kan schemaläggas på andra tider.

### Avbrott och avanmälan på kurs

Enligt beslut vid Linköpings universitet om Riktlinjer och rutiner för bekräftelse av deltagande i utbildning med mera på grund- och avancerad nivå, Dnr LiU-2020-02256 (<https://styrdokument.liu.se/Regelsamling/VisaBeslut/764582>) skall avbrott i studier registreras i Ladok. Alla studenter som inte deltar i kurs man registrerat sig på är alltså skyldiga att anmäla avbrottet så att detta kan noteras i Ladok. Avanmälan eller avbrott från kurs görs via webbformulär [Blanketter och formulär](#)

### Inställd kurs eller avvikelse från kursplanen

Kurser med få deltagare (< 10) kan ställas in eller organiseras på annat sätt än vad som är angivet i kursplanen. Om kurs skall ställas in eller avvikelse från kursplanen skall ske prövas och beslutas detta av dekanus. För fristående kurser måste inställande av kurs ske innan studenter har antagits på kursen (i enlighet med LiUs antagningsordning Dnr LiU-2022-01200, <https://styrdokument.liu.se/Regelsamling/VisaBeslut/622645>).

### Riktlinjer rörande examination och examinator

Se Beslut om Riktlinjer för utbildning och examination på grundnivå och avancerad nivå vid Linköpings universitet Dnr LiU-2023-00379, (<http://styrdokument.liu.se/Regelsamling/VisaBeslut/917592>).

Examinator för en kurs ska inneha en läraranställning vid LiU i enlighet med LiUs anställningsordning, Dnr LiU-2022-04445 (<https://styrdokument.liu.se/Regelsamling/VisaBeslut/622784>). För kurser på avancerad nivå kan följande lärare vara examinator: professor (även adjungerad och gästprofessor), biträdande professor (även adjungerad), universitetslektor (även adjungerad och gästlektor), biträdande universitetslektor eller postdoktor. För kurser på grundnivå kan följande lärare vara examinator: professor (även adjungerad och gästprofessor), biträdande professor (även adjungerad), universitetslektor (även adjungerad och gästlektor), biträdande universitetslektor, universitetsadjunkt (även adjungerad och gästadjunkt) eller

postdoktor. I undantagsfall kan även en Timlärare utses som examinator på både grund- och avancerad nivå, se Tekniska fakultetsstyrelsen vidaredelegationer.

## Examination

### Principer för tentamina

Skriftlig och muntlig tentamen samt digital salstentamen och datortentamen ges minst tre gånger årligen; en gång omedelbart efter kursens slut, en gång i augustiperioden samt vanligtvis i en av omtentamensperioderna. Annan placering beslutas av programnämnden.

Principer för tentamensschemat för kurser som följer läsperioderna:

- kurser som ges Vt1 förstagångstenteras i mars och omtenteras i juni och i augusti
- kurser som ges Vt2 förstagångstenteras i maj och omtenteras i augusti och i januari
- kurser som ges Ht1 förstagångstenteras i oktober och omtenteras i januari och augusti
- kurser som ges Ht2 förstagångstenteras i januari och omtenteras i mars och i augusti

Tentamensschemat utgår från blockindelningen men avvikelser kan förekomma främst för kurser som samläses/samtenteras av flera program samt i lägre årskurs.

För kurser som av programnämnden beslutats vara vartannatårskurser ges tentamina 3 gånger endast under det år kursen ges.

För kurser som flyttas eller ställs in så att de ej ges under något eller några år ges tentamina 3 gånger under det närmast följande året med tentamenstillfällen motsvarande dem som gällde före flyttningen och/eller inställandet av kursen.

När en kurs, eller ett tentamensmoment (TEN, DIT, DAT, MUN), ges för sista gången ska ordinarie tentamen och två omtentamina erbjudas. Därefter fasas examinationen ut under en avvecklingsperiod med tre tentamina samtidigt som tentamen ges i eventuell ersättningskurs under det följande läsåret. Undantaget är kurser som gavs i perioden HT1, där de tre examinationstillfällena blir januari, mars och augusti. Om ingen ersättningskurs finns ges tre tentamina i omtentamensperioder under det följande läsåret. Annan placering beslutas av programnämnden. I samtliga fall ges dessutom tentamen ytterligare en gång under det därpå följande året om inte programnämnden föreskriver annat. Totalt erbjuds alltså 6 omtentamenstillfällen, varav 2 ordinarie omtentamenstillfällen. I tentaansmälningssystemet markeras tentamina som ges för näst sista respektive sista gången.

Om en kurs ges i flera perioder under året (för program eller vid skilda tillfällen för olika program) beslutar programnämnden/programnämnderna gemensamt om placeringen av och antalet omtentamina.

För fristående kurser med tentamensmoment som inte följer blockplacering kan

andra tider förekomma.

### **Omprov övriga examinerande moment**

För riktlinjer för omprov vid andra examinerande moment än skriftliga tentamina, digital salstentamina och datortentamina hänvisas till de generella LiU-riktlinjerna för examination och examinator, Dnr LiU-2023-00379 (<http://styrdokument.liu.se/Regelsamling/VisaBeslut/917592>).

### **Nedlagd kurs**

För Beslut om Rutiner för administration vid avveckling av utbildningsprogram, fristående kurser och kurser inom program, se Dnr LiU-2021-04782 (<https://styrdokument.liu.se/Regelsamling/VisaBeslut/1156410>). Efter beslut om nedläggning och efter avvecklingsperiodens slut hänvisas studenterna till ersättande kurs (eller motsvarande) enligt information i kursplan eller utbildningsplan. Om en student har godkänt i något/några moment i en avvecklad programkurs men inte alla och det finns en åtminstone delvis ersättande kurs så kan en bedömning om eventuellt tillgodoräknande ske. Vid eventuella frågor om tillgodoräkning av del av kurs kontakta studievägledare.

### **Anmälan till tentamen**

För deltagande i skriftlig tentamen, digital salstentamen och datortentamen är anmälan obligatorisk, se beslut i regelsamlingen Dnr LiU-2020-04559 (<https://styrdokument.liu.se/Regelsamling/VisaBeslut/622682>). En oanmäld student kan således *inte* erbjudas plats. Anmälan till tentamen är öppen 30 kalenderdagar före provdatum och stänger 10 kalenderdagar innan provdatum om inget annat anges. Anmälan görs i Studentportalen eller via LiU-appen. Anvisad sal meddelas fyra dagar före tentamensdagen via e-post.

### **Ordningsföreskrifter för studerande vid tentamensskrivningar**

Se särskilt beslut i regelsamlingen, Dnr LiU-2020-04559 (<http://styrdokument.liu.se/Regelsamling/VisaBeslut/622682>).

### **Plussning**

Vid Tekniska högskolan vid LiU har studerande rätt att genomgå förnyad examination (s.k. plussning) för högre betyg på skriftliga tentamina, digital salstentamina och datortentamina, dvs samtliga provmoment med modulkod TEN, DIT och DAT. På övriga examinationsmoment ges inte möjlighet till plussning, om inget annat anges i kursplan.

Plussning är ej möjlig på kurser som ingår i utfärdad examen.

### **Betyg och examinationsformer**

Företrädesvis skall betygen underkänd (U), godkänd (3), icke utan beröm godkänd (4) och med beröm godkänd (5) användas.

- Kurser med skriftlig tentamen och digital salstentamen skall ge betygen (U, 3, 4, 5).

- Kurser med stor del tillämpningsinriktade moment såsom laborationer, projekt eller grupparbeten får ges betygen underkänd (U) eller godkänd (G).
- Examensarbete samt självständigt arbete ger betyg underkänd (U) eller godkänd (G).

### Examinationsmoment och modulkoder

Nedan anges vad som gäller för de examinationsmoment med tillhörande modulkod som tillämpas vid Tekniska fakulteten vid Linköpings universitet.

- Skriftlig tentamen (TEN) och digital salstentamen (DIT) skall ge betyg (U, 3, 4, 5).
- Examinationsmoment som kan ge betygen underkänd (U) eller godkänd (G) är laboration (LAB), projekt (PRA), kontrollskrivning (KTR), digital kontrollskrivning (DIK), muntlig tentamen (MUN), datortentamen i datorsal (DAT), uppgift (UPG), hemtentamen (HEM), digital kontrollskrivning i datorsal (DAK).
- Övriga examinationsmoment där examinationen uppfylls framför allt genom aktivt deltagande som basgrupp (BAS) eller moment (MOM) ger betygen underkänd (U) eller godkänd (G).
- Examinationsmomenten Opposition (OPPO) och Auskultation (AUSK) inom examensarbetet ger betyg underkänd (U) eller godkänd (G).

Allmänt gäller att:

- Obligatoriska kursmoment skall vara poängsatta och ges en modulkod.
- Examinationsmoment som ej är poängsatt får ej vara obligatoriskt. Det är frivilligt att delta på dessa moment och information om det samt tillhörande villkor skall tydligt framgå i den beskrivande texten.
- För kurser med flera examinationsmoment med graderad betygsskala skall det anges hur slutbetyg på kursen vägs samman.

För obligatoriska moment gäller att (i enlighet med Riktlinjer för utbildning och examination på grundnivå och avancerad nivå vid Linköpings universitet, Dnr LiU-2023-00379 <http://styrdokument.liu.se/Regelsamling/VisaBeslut/917592>):

- Om det finns särskilda skäl, och om det med hänsyn till det obligatoriska momentets karaktär är möjligt, får examinator besluta att ersätta det obligatoriska momentet med en annan likvärdig uppgift.

För möjlighet till anpassade examinationsmoment gäller att (i enlighet med Riktlinjer för utbildning och examination på grundnivå och avancerad nivå vid Linköpings universitet, Dnr LiU-2023-00379 <http://styrdokument.liu.se/Regelsamling/VisaBeslut/917592>):

- Om LiU:s koordinator för studenter med funktionsnedsättning har beviljat en student rätt till anpassad examination vid salstentamen har studenten rätt till det.
- Om koordinatören har gett studenten en rekommendation om anpassad examination eller alternativ examinationsform, får examinator besluta om



detta om examinator bedömer det möjligt utifrån kursens mål.

- Examinator får också besluta om anpassad examination eller alternativ examinationsform om examinator bedömer att det finns synnerliga skäl och examinator bedömer det möjligt utifrån kursens mål.

### **Rapportering av examinationsresultat**

Rapportering av den studerandes examinationsresultat sker på respektive institution.

### **Plagiering**

Vid examination som innebär rapportskrivande och där studenten kan antas ha tillgång till andras källor (exempelvis vid självständiga arbeten, uppsatser etc) måste inlämnat material utformas i enlighet med god sed för källhänvisning (referenser eller citat med angivande av källa) vad gäller användning av andras text, bilder, idéer, data etc. Det ska även framgå ifall författaren återbrukat egen text, bilder, idéer, data etc från tidigare genomförd examination, exempelvis från kandidatarbete, projektrapporter etc. (ibland kallat självplagiering).

Underlåtelse att ange sådana källor kan betraktas som försök till vilseledande vid examination.

### **Försök till vilseledande**

Vid grundad misstanke om att en student försökt vilseleda vid examination eller när en studieprestation ska bedömas ska enligt Högskoleförordningens 10 kapitel examinator anmäla det vidare till universitetets disciplinnämnd. Möjliga konsekvenser för den studerande är en avstängning från studierna eller en varning. För mer information se [Fusk och plagiat](#).

Linköpings universitet har även tagit fram en vägledning för lärares och studenters användning av generativ AI i utbildningen (Dnr LiU-2023-02660). Som student förväntas du alltid ta reda på vad som gäller för respektive kurs (inklusive examensarbetet). Generellt gäller tydlighet för var och hur generativ AI har använts.

### **Regler**

Universitetet är en statlig myndighet vars verksamhet regleras av lagar och förordningar, exempelvis Högskolelagen och Högskoleförordningen. Förutom lagar och förordningar styrs verksamheten av ett antal styrdokument. I Linköpings universitets egna regelverk samlas gällande beslut av regelkaraktär som fattats av universitetsstyrelse, rektor samt fakultets- och områdesstyrelser.

LiU:s regelsamling angående utbildning på grund- och avancerad nivå nås på <https://styrdokument.liu.se/Regelsamling/Innehall>.