

# Runskrift



ECUTBILDNING

Anton Grigoriev

EC Utbildning

2024–10

# Abstract

This project applies machine learning to analyze and classify runic inscriptions. The data is then integrated into Power BI to create interactive visualizations, including a map displaying the geographical distribution of runestones. By leveraging these technologies, we gain valuable insights into the history and patterns of runestone inscriptions.

## Innehållsförteckning

Abstract .....	2
1 Inledning .....	1
1.1 Runinskrifter .....	1
1.2 Runinskrifters datering .....	1
2 Syfte .....	2
3 Teori.....	2
3.1 Word2Vec .....	2
3.2 Naive Bayes.....	3
3.3 KNN.....	3
4 Metod .....	3
4.1 Datainsamling .....	3
4.2 Data modellering. ....	4
4.3 Agil arbetsmetodik.....	4
4.3.1 GitHub.....	5
5 Resultat och Diskussion .....	5
6 Slutsatser .....	7
7 Självutvärdering.....	7
Källförteckning .....	7

# 1 Inledning

## 1.1 Runinskrifter

Runinskrifter är viktigt källa i språk och historieforskning. Dem flesta bevarade runinskrifter finns på runstenar. Ungefär 800 medeltida inskrifter bevarades på trädskivor i Bergen, Norge, åtskilligt mycket hittas vid arkeologiska utgrävningar på mynt, keramik eller metallbleck. Den sistnämnda innehåller ofta ett eller några få runor och går inte att förstå. Alla läsbara runinskrifter samlas och publiceras med signum, t.ex. U 1072 blir runinskrift från Uppland (U) med nummer 1072.

Förr i tiden hade runstenar flera olika funktioner:

minnessten över den döda

arvsdokument

demonstration av status och rikedom

andra funktioner, som vägvisare, gränsmärke o.s.v.

Vikingatida kvinnor kunde ärva efter sina makar och barn, i sådana komplicerade förhållanden skrevs ibland långa inskriptioner på fasta hällar.

Minnesstenarna ristades enligt enkelt minnesformula: NN reste stenen efter YY. Ibland ligger man till några beskrivande ord om YY eller NN. Sedan lagdes ristarens signatur, som hade en del av notarie eller bestyrkarefunktion.

Geografisk placering av stenarna, alltså var och hur nära placerades stenarna, hjälper att bestämma om samma namn på olika stenar avser samma person, i vilken region vistades samma ristare.

Spridning över landet kan hjälpa att studera forntida dialekter och förstå mer om struktur av vikingatida samhället.

## 1.2 Runinskrifters datering

De flest av de läsbara runinskrifter publiceras utan årtal. Man känner helt enkelt inte när inskrifterna ristades. Samnordisk runtextdatabas,<sup>[1]</sup> som samlar inskrifterna har flesta av de grovt indelade i klasser

- U: urnordiska
  - M: medeltida
  - Sentida
- Saknas klassen, då är inskriften V: vikingatida.

Flesta dateringarna repeterar klassen, ibland med osäkerhet (V/M eller V?).

Möjlighet till datering erbjöds av språk- och skriftegenskaper, tidssekvenser<sup>[2]</sup> av minnesinskrifter mellan samma familj (t.ex. son efter fader) eller anhöriga, stildrag eller runstilar<sup>[3]</sup>:

**RAK****Fp**

Unornamented stones:	c. 980? – 1015
Bird's-eye-view:	c. 1010 – 1050
Pr 1:	c. 1010 – 1040
Pr 2:	c. 1020 – 1050
Pr 3:	c. 1045 – 1075
Pr 4:	c. 1070 – 1100
Pr 5:	c. 1100 – 1130

Tabell 1: runstilar med dateringar.<sup>[3]</sup>

U 1065 på bild och i verklighet.

Serier eller grupper av ristningar presenteras vanligtvis på kartan och man kan t.ex. följa enskilda ristaren i hans arbetsresor<sup>[4]</sup>.

Förekommande metod för datering är att jämföra stil med andra daterbara stildrag efter föremålet/inskriften, som t.ex. korsform.<sup>[2]</sup> P.g.a. många runstenar är täckta med lavar och äldre avbildningar är inte alltid pålitliga är denna metod problematiskt för massdatering.

Dateringar av runstilar har stor överlapp (ristningarna utan dekoration (RAK) förekommer nästan alltid) och är snarare ett förslag eller hypotes.

## 2 Syfte

Syftet med denna rapport är försöka datera eller klassificera ristningarna i tidsperioder. För att uppfylla syftet så kommer följande frågeställning(ar) att besvaras:

- 1) Finns det betydande skillnader mellan runinskrifterna för att tillåta klassificering i olika tidsperioder?
- 2) Är det möjligt att slutföra dateringen av runstenarna med en annan strategi än den visuella observationen av individuella stildrag eller komplement till den?
- 3) Är det möjligt att kontrollera dateringen genom kartläggning av olika ristningar och deras geografiska utbredning?

## 3 Teori

### 3.1 Word2Vec

För att möjliggöra arbete med runtext i translittererad form word2vec modell tränades på alla ristningar utan någon kännedom till andra språk eller vokabulär. P.g.a. minnesformula är "NN reste stenen efter YY", alla ord inklusive partiklar och prepositioner användes, för att även partikel kan betydande ändra formeln.

word2vec förstår inte enstaka ord, utan behåller kontext för varje ord i vektorform. Vi sparar ristningarna i vektorform genom att summera kontexter för alla ord. Vissa enstaka ord eller namn som sällan användes har ingen statistiskt signifikanta kontext och lämnas ifrån vokabulären av ord kontexter, men lämnar sitt spår i kontexter av andra ord.<sup>[5]</sup>

Exempel:

```

# one can really have fun with Word2Vec – give first 10 most important words from context of stain :
print(Word2Vec_model.wv.most_similar(positive=['stain'], topn=10))
[('auk', 0.999559819698337), ('at', 0.9994414448738098), ('sin', 0.9993997812271118), ('stan',
0.9993939995765686), ('stin', 0.9993836879730225), ('bru', 0.9993816614151001), ('a',
0.9993793368339539), ('sun', 0.9993706941604614), ('litu', 0.9993685483932495), ('til',
0.9993651509284973)]
# representation of a word in 100D space :
Word2Vec_model.wv['stain']
array([ 0.01777048, -0.2663258 , -0.08850657, -0.21156931,  0.3107939 ,
       0.17151207, -0.07560526, -0.46269372,  0.14711604,  0.15908192,
      0.22800867,  0.00301661,  0.27782336,  0.40926263,  0.23630734,
      0.03280863,  0.09002124, -0.2586201 , -0.24972142,  0.24430792,
      ...
     -0.499536 ,  0.06716365, -0.16504884,  0.14478783,  0.27548143,
      0.42472345,  0.00899772, -0.14505339, -0.07245647,  0.16830586,
     -0.03517922,  0.00225671, -0.1718785 , -0.66785014, -0.11197505,
      0.02186924, -0.01297164,  0.15166874, -0.36323604,  0.22469532,
     -0.09646068,  0.02419424,  0.25739014, -0.3346674 , -0.01831245,
      0.3605978 ,  0.03796249, -0.01960598, -0.14870098,  0.20199582],
      dtype=float32)

```

### 3.2 Naive Bayes

Naive Bayes utgår ifrån att variablerna är oberoende av varandra och att alla variablerna bidrar lika mycket till resultatet. Detta antagande stämmer inte för de flesta verkliga situationer, men det förenklar de numeriska beräkningarna och gör problemet lättare hanteringsbart.<sup>[7]</sup>

År 2013 användes redan algoritmen för att datera runstenar fast utan att använda ordförråd eller själva inskrifterna.<sup>[6]</sup> Försöket då hade ingen stor succé. Jag ville prova med samma metod och med word2vec modell för ristningarna men kunde inte hitta fungerande NB variant för regression. Klassificeringsvariant användas dock utan ristningstexter.

### 3.3 KNN

Med K-Nearest Neighbors (KNN) lagrar man alla tillgängliga värden och klassificerar varje ny data baserat på dess likhet med datapunkten som ligger närmast. Tanken bakom modellen är att punkter som ligger nära varandra brukar tillhöra samma kategori.<sup>[8]</sup>

KNN anses ibland att bli näst enkelt efter Naive Bayes och är känd för att fungera bra med word2vec. Den används i beräkningarna med default sättningarna.<sup>[5]</sup>

## 4 Metod

### 4.1 Datainsamling

Vi samlade data från två olika källor för att möjliggöra jämförelse: Samnordisk runtextdatabas finns tillgängligt gratis på Uppsala universitetens webbsida och Wikipedia användes som alternativ.

Wikipedias API används både direkt via web interface med "requests" modul och med "wikipediaapi" wrapper. Den första kan access wikipedias mallar med data, den andra kan lösa problem om runinskrift är angiven utan mallen. Med både runtexts översättning och annan text som vanligt svenskt text gör det svårt att urskilja just översättning. Kan man inte hitta 'Runskriftsöversättning' i

mallen måste man söka för texten mellan '<p>Översättning\xA0till nusvenska:\n</p>\n<dl><dd>' och '</dd></dl>' i vanlig html.

Data rensades med filter i python som

```
raw_of_data = [ele.replace('&nbsp;', ' ') if isinstance(ele,str) else ele for ele in raw_of_data] #replace hard spaces
och sedan för hand i .csv textfil.
```

translitterering	normalisering	translation	Latitude	Longitude	Material	Period_begin	Period_end	...	Östen 2	puliR	Fp	KB	RAK	...
kinluk x hulmkis x tutir x systir x sukruþar x...	Ginnlaug, Holmgæirs dottir, systir Sygröðar ok... Ingi... ok logærðr þau letu íuker [:] þau : itu : rais...	Ginnlög, Holmgers dotter, syster till Sygröd o... Inge--- och Jorgard de läto räisa stæin þenna ... resa denna sten ef...	59.508240	17.627210	runsten	793	1066	...	0.0	0.0	0.0	0.0	0.0	0.0
· inki-... [ajuk : iuker [:] þau : itu : rais...	Ingi... ok logærðr þau letu íuker [:] þau : itu : rais...	Inge--- och Jorgard de läto räisa stæin þenna ... resa denna sten ef...	59.768910	17.412310	runsten	1000	1066	...	0.0	0.0	0.0	0.0	0.0	0.0

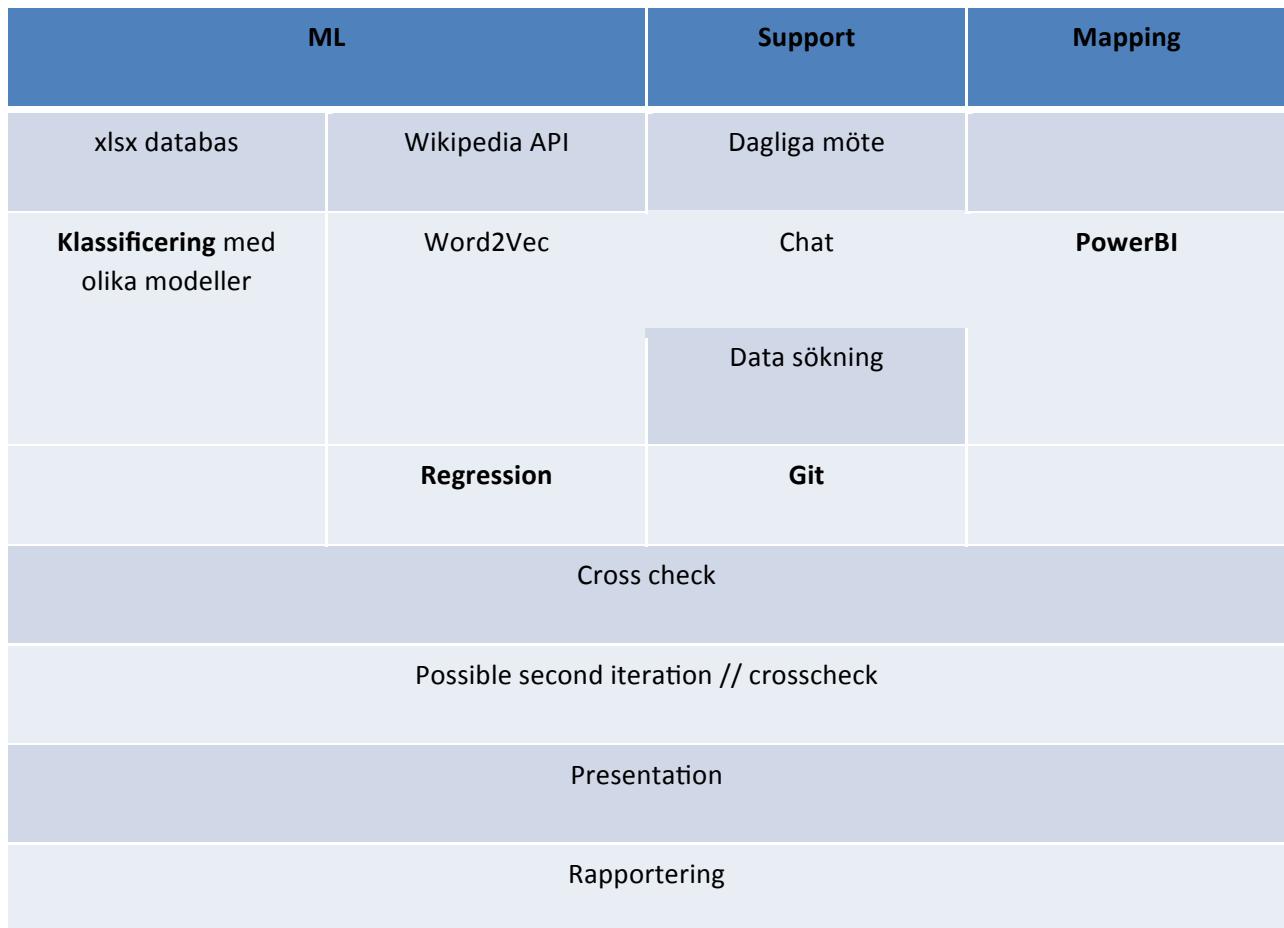
Typisk data från Wikipedia, Östen är ristare, Fp, KB och RAK är stilar. Notera att KB är inte med i Samnordisk runtextdatabas.

## 4.2 Data modellering.

KNN tränades på ~40 stenar daterade med runstenstilsmetod. Normalisering gjordes förhand. Data för Word2vec sammansättes från word2vec, stil, ristare, koordinater och period.

## 4.3 Agil arbetsmetodik

Vi skapade arbetsplan och delade sig i tre grupper för att möjliggöra effektivt samverkning.



Planet utgår ifrån att samtidigt skapa datering av ristningar och förbereda presentation på kartan av daterade stenar eller sekvenser av stenar, t.ex. från samma ristare.

Vi höll vardagliga möten och följe chatten i Teams.

### 4.3.1 GitHub

I vår GitHub-repo hittar du en omfattande README-fil som fungerar som guide till projektet. Vi börjar med att ge en tydlig överblick över vad projektet syftar till.

Steg för steg leder vi dig genom projektets arbetsflöde, så att du enkelt kan förstå hur alla delar hänger ihop. Vi har också tagit fram en detaljerad installationsguide som tydliggör steg för steg installationsprocessen. Här hittar du all nödvändig information om vilka programvaror och bibliotek du behöver

Med README-filen som din guide kommer du snabbt att kunna komma i gång med att använda och utforska projektet på egen hand. Några av fördelarna med GitHub är att det blir lättare att samarbeta med flera olika utvecklare som arbetar på samma projekt. GitHub gör det enkelt att se vem som har gjort vilka ändringar och att slå samman olika versioner av kod.

GitHub har en Project board feature där vi skapade en Kanban för att kunna spåra framsteg på ett snabbt sätt. Där organiserade vi olika uppgifter och funktioner i tre olika stadier som "In Progress", "Ready" och "Done", så att gruppen kunde visualisera projektets övergripande status.

## 5 Resultat och Diskussion

Man kan förbättra klassificering av ristningar i Samnordisk runtextdatabas, men också datera alla ristningar med regression. Jag visar med exempel på serie av ristningar från ristare Tidkumle:

```
# test on inscribers like Tidkume (person)

mask = df['Inscriber'].str.contains('Tidkume', na=False, regex=False) #Öpir Tidkume

X_pred = df_ML[mask]

X_pred = X_pred.fillna(0)

X_pred.columns = X_pred.columns.astype(str)

y_pred = neigh.predict(X_pred)

print(y_pred)

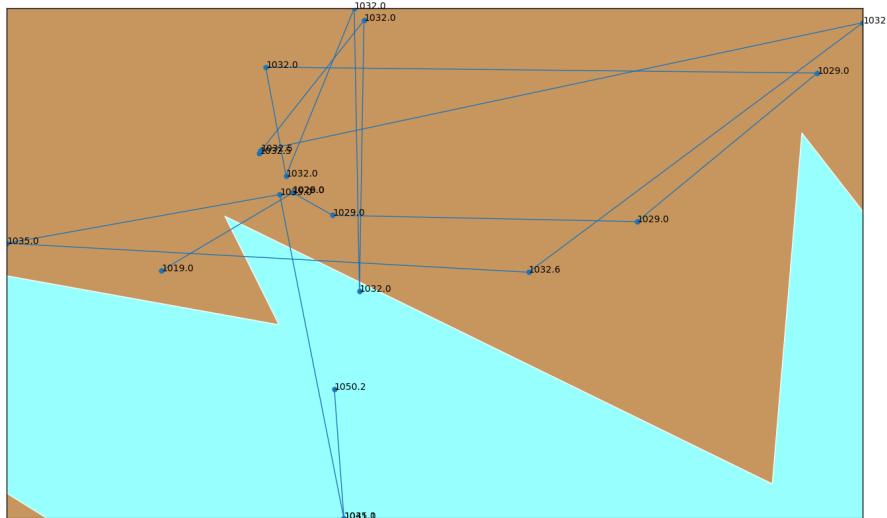
[1050.2 1032.6 1032. 1029. 1029. 1032.5 1032.5 1026. 1029. 1035.

1032. 1019. 1035. 1032. 1032. 1029. 1032.5 1032. 1041.1 1035.]
```

På nästa sidan kan man se att de två första årtal rör stenar med osäkra attribution, så att Tidkumle verkade från 109 till 1035 – rimligt tid. Han har bara en stil markerad för hans inskrifter och stenar ristade av honom nära varandra har olika dateringar med andra stenar långt borta mittemellan. Det innebär att plats och stil inte har direkta påverkan på dateringen, utan utgör den snarare från textens analys möjligent indirekt kopplad till stil och plats information.

			signum	revisionID	translitterering	normalisering	translation	Latitude	Longitude	Stil	Inscriber
606	Upplands runinskrifter 691	54589581	birn : lit : risa : s--- ...- : iftir : sterbi...	Bjorn let ræisa s[tæin þenna] æftir Styrbiorn...	Björn lät resa (denna sten) efter Styrbjörn,	59.49639	17.12510	Nan	Ödbjörn enligt Erik Brate,Brate, ej troligt, ...		
621	Upplands runinskrifter 716	54542996	· bruni : lit : resa : auk : ari stin : þina ....	Bruni let ræisa ok Ari()/Arni() stæin pennna æ...	Brune lät resa och Are() denna sten efter Gudf...	59.56626	17.24156	Nan	Ödbjörn enligt Erik Brate,Brate. Osannolikt e...		
622	Upplands runinskrifter 718	54593741	[...asr : auk : olifr : auk : uar:asi : li-.....]	... ok Olæifr ok Var-Asi()/Varr-Asi() le[tu] ....	... och Olev och ""uarasi"" läto ... denna s...	59.55498	17.14018	Nan	Tidkume , osäker bedömning		
635	Upplands runinskrifter 733	54537884	[(þ)(i)(l)stan : lit : ris... ... .n : sun ....]	Þorstæinn() let ræis[a] ... ..., sun sinn goða...	Torsten() lät resa ..., sin gode son. Tidkume ...	59.59654	17.30669	Nan	Tidkume		

	Latitude	Longitude	0	signum
11	59.56718	17.02141	1019.0	Upplands runinskrifter 775
7	59.61380	17.10020	1026.0	Upplands runinskrifter 763
8	59.61386	17.10017	1029.0	Upplands runinskrifter 764
4	59.60033	17.12405	1029.0	Upplands runinskrifter 742
3	59.59654	17.30669	1029.0	Upplands runinskrifter 733
15	59.68542	17.41468	1029.0	Upplands runinskrifter 813
10	59.68896	17.08397	1032.0	Upplands runinskrifter 771
17	59.62408	17.09613	1032.0	Upplands runinskrifter ATA6243/65
14	59.72419	17.13704	1032.0	Upplands runinskrifter 803
2	59.55498	17.14018	1032.0	Upplands runinskrifter 718
13	59.71691	17.14289	1032.0	Upplands runinskrifter 800
5	59.63749	17.08005	1032.5	Upplands runinskrifter 758
6	59.63898	17.08096	1032.5	Upplands runinskrifter 759
16	59.71556	17.44202	1032.5	Upplands runinskrifter 828
1	59.56626	17.24156	1032.6	Upplands runinskrifter 716
12	59.58356	16.92860	1035.0	Upplands runinskrifter 777
9	59.61276	17.09235	1035.0	Upplands runinskrifter 765
19	59.41897	17.13084	1035.0	Södermanlands runinskrifter 377
18	59.41912	17.13087	1041.1	Södermanlands runinskrifter 205
0	59.49639	17.12510	1050.2	Upplands runinskrifter 691



Tidkumles resor med predikterad tid och känd plats markerade.

Tider för andra kända ristningsserier som t.ex. ristningar från Öpirs verksamhet, Ingvarståg, tegnstenar,<sup>[9]</sup> kommer också med rimligt tidsspann.

## 6 Slutsatser

Man kan datera runinskrifter både med Samnordisk runtextdatabas utan text och med "bara" text. Ska leda till synergi!

Man kan jämföra med andra kända försök att datera t.ex. Öpirs ristningar<sup>[2]</sup> eller tegnstenar. Med ungefär  $\pm 15$  år korrekt träningsdata får man ungefär  $\pm 20$  till ungefär  $\pm 30$  års serier av helt meningsfulla prediktioner.

Bättre modell än KNN, bättre data än wikipedia, men framförallt synergi med stilar och andra data från Samnordisk runtextdatabas kan leda till ännu bättre resultat.

## 7 Självutvärdering

1. Utmaningar du haft under arbetet samt hur du hanterat dem.  
Att rensa data var den största utmaningen under projektet, samt jobba med API. Man ska ha tålmod.
2. Vilket betyg du anser att du skall ha och varför.  
VG
3. Något du vill lyfta fram till Antonio?  
nej

## Källförteckning

1. ["Samnordisk runtextdatabas - Uppsala universitet"](https://www.uu.se/institution/nordiska/forskning/projekt/samnordisk-runtextdatabas). [www.uu.se](http://www.uu.se). 20 maj 2024.  
<https://www.uu.se/institution/nordiska/forskning/projekt/samnordisk-runtextdatabas>. Läst 1 november 2024.
2. Tesch, Sten. ["Uppländska runstenar och det långa skiftet: nya tolkningsmöjligheter."](#). SKIFTET – vikingatida sed och kristen tro. Ett mångvetenskapligt perspektiv på kristnandeprocessen i Mälardalen. Artos förlag, Skellefteå. 2017. FULL TEXT.

- [https://www.academia.edu/36259060/SKIFTET\\_vikingatida\\_sed\\_och\\_kristen\\_tro\\_Ett\\_m%C3%A5ngvetenskapligt\\_perspektiv\\_p%C3%A5\\_kristnandprocessen\\_i\\_M%C3%A4laromr%C3%A5det\\_Artos\\_f%C3%B6rlag\\_Skellefte%C3%A5\\_2017\\_FULL\\_TEXT](https://www.academia.edu/36259060/SKIFTET_vikingatida_sed_och_kristen_tro_Ett_m%C3%A5ngvetenskapligt_perspektiv_p%C3%A5_kristnandprocessen_i_M%C3%A4laromr%C3%A5det_Artos_f%C3%B6rlag_Skellefte%C3%A5_2017_FULL_TEXT). Läst 1 november 2024.
3. Gräslund, Anne-Sofie (2006-01-01). ["Dating the Swedish Viking-Age rune stones on stylistic grounds"](https://www.academia.edu/6107509/Dating_the_Swedish_Viking_Age_rune_stones_on_stylistic_grounds). *Runes and their secrets: studies in runology*, ed. by M. Stoklund, M. Lerche Nielsen, B. Holmberg & G. Fellows-Jensen. Copenhagen: Museum Tusculanum Press, 2006en. [https://www.academia.edu/6107509/Dating\\_the\\_Swedish\\_Viking\\_Age\\_rune\\_stones\\_on\\_stylistic\\_grounds](https://www.academia.edu/6107509/Dating_the_Swedish_Viking_Age_rune_stones_on_stylistic_grounds). Läst 1 november 2024.
  4. Kitzler Åhfeldt, Laila (2022). *Runristare och mobilitet i Södermanland*. Riksantikvarieämbetet. [ISBN 978-91-7209-807-7](#). Läst 1 november 2024
  5. ["How to classify text using Word2Vec - Thinking Neuron"](https://thinkingneuron.com/how-to-classify-text-using-word2vec/) (på amerikansk engelska). 15 december 2021. <https://thinkingneuron.com/how-to-classify-text-using-word2vec/>. Läst 1 november 2024.
  6. Parth Shukla (4 januari 2023). ["Naive Bayes Algorithms: A Complete Guide for Beginners"](https://www.analyticsvidhya.com/blog/2023/01/naive-bayes-algorithms-a-complete-guide-for-beginners/) (på engelska). *Analytics Vidhya*. <https://www.analyticsvidhya.com/blog/2023/01/naive-bayes-algorithms-a-complete-guide-for-beginners/>. Läst 1 november 2024.
  7. Thoeming, Alix (2013-01-01). ["Exploring interconnectivity and similarity in the rune-stones of 10th-12th century Sweden"](https://www.academia.edu/88876974/Exploring_interconnectivity_and_similarity_in_the_rune_stones_of_10th_12th_century_Sweden). [https://www.academia.edu/88876974/Exploring\\_interconnectivity\\_and\\_similarity\\_in\\_the\\_rune\\_stones\\_of\\_10th\\_12th\\_century\\_Sweden](https://www.academia.edu/88876974/Exploring_interconnectivity_and_similarity_in_the_rune_stones_of_10th_12th_century_Sweden). Läst 1 november 2024.
  8. Ali Omer Bajallan, Rebwar (2022). ["A comparative evaluation of machine learning models for engagement classification during presentations : A comparison of distance- and non-distance-based machine learning models for presentation classification and class likelihood estimation."](https://www.diva-portal.org/smash/record.jsf?pid=diva2:1732944&dswid=6146) <https://www.diva-portal.org/smash/record.jsf?pid=diva2:1732944&dswid=6146>. Läst 1 november 2024
  9. Hagerman, Maja (1999). *Spåren av kungens män: om när Sverige blev ett kristet rike i skiftet mellan vikingatid och medeltid* (2. uppl). Prisma. [ISBN 978-91-518-3694-2](#). Läst 1 november 2024