

Hagnýt notkun gervigreindar í rannsóknum *og takmörkun þess*

Halldór Bjarki Ólafsson

BSc Sameindalíffræði, Læknir, Doktorsnemi í læknavísindum



Halldór Bjarki Ólafsson

BSc Sameindalíffræði 2016



BSc Læknisfræði 2019



Kandídatspróf Læknisfræði 2022



Doktorsnám og rannsóknir



BJA

British Journal of Anaesthesia, 124 (6): 718–725 (2020)
doi: 10.1016/j.bja.2020.02.009
Advance Access Publication Date: 23 March 2020
Clinical Practice

A retrospective cohort study on the association between elevated preoperative red cell distribution width and all-cause mortality after noncardiac surgery

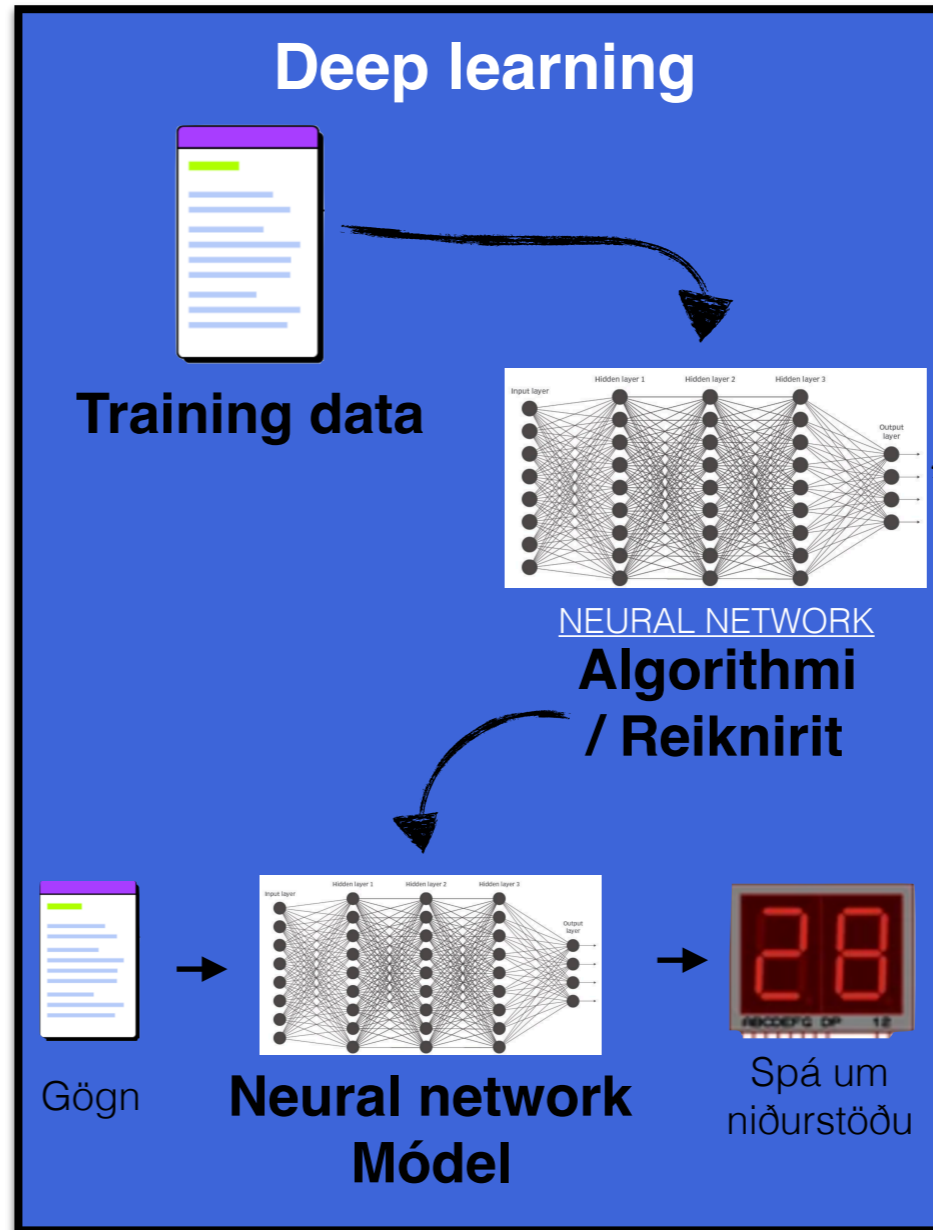
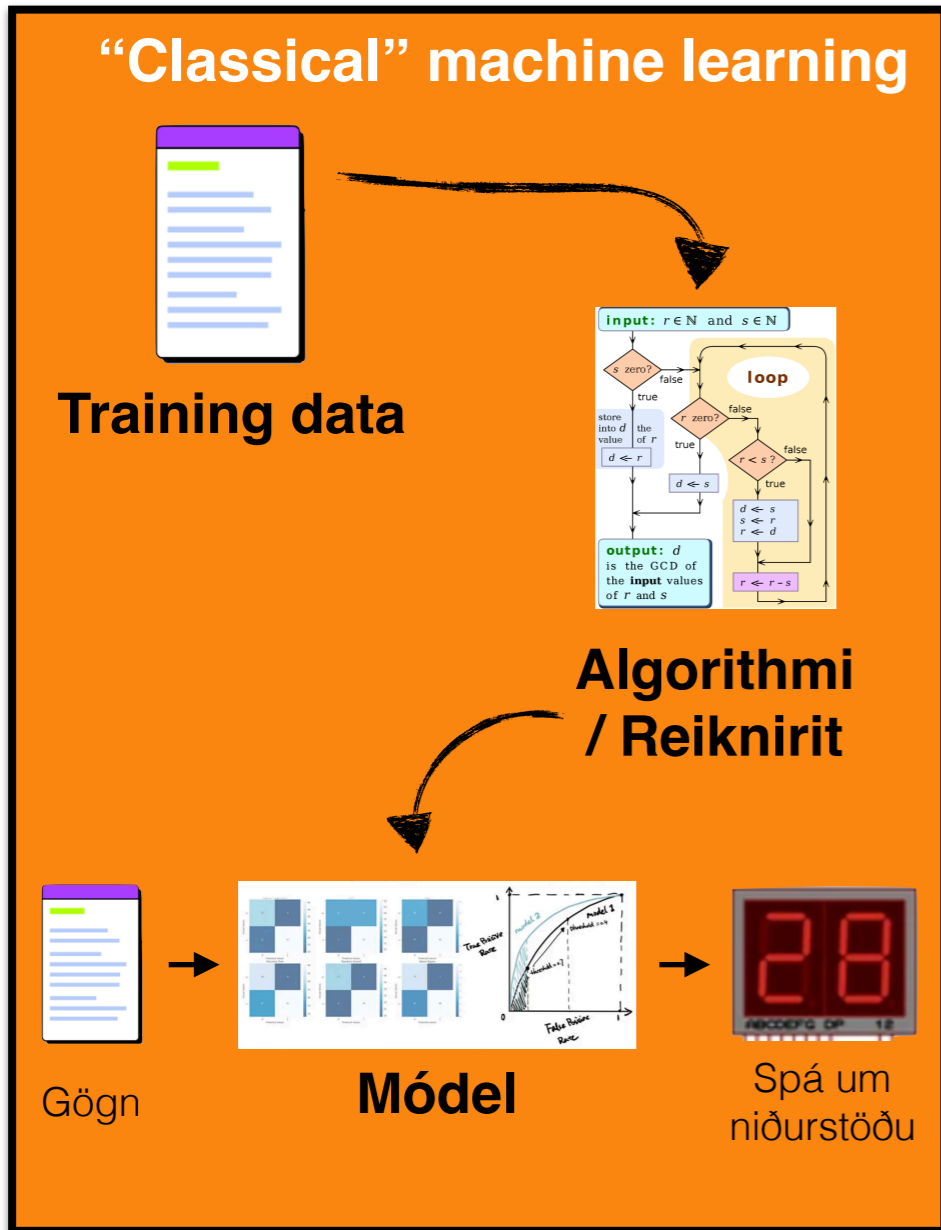
Halldor B. Olafsson¹, Gissur A. Sigurdarson¹, Kenneth B. Christopher^{2,3}, Sigurbergur Karason^{1,4}, Gísli H. Sigurdsson^{1,4} and Martin I. Sigurdsson^{1,4,*}

¹Faculty of Medicine, University of Iceland, Reykjavik, Iceland, ²Division of Renal Medicine, Brigham and Women's Hospital, Boston, MA, USA, ³Harvard Medical School, Boston, MA, USA and ⁴Division of Anaesthesia and Intensive Care Medicine, Landspítali, National University Hospital of Iceland, Reykjavik, Iceland

*Corresponding author. E-mail: martinigi@gmail.com



- **Hluti 1: Hvað er gervigreind?**
 - Classical machine learning og deep learning
 - ChatGPT og verkunarmáti
- **Hluti 2: Rannsóknir læknávísinda**
 - Faraldsfræðilegar rannsóknir
 - Hraðatakmarkandi skref í rannsóknum
 - Hagnýting gervigreindar og takmarkanir
- **Hluti 3: Samantekt** - Umræður og spurningar.



GEOFFREY E HINTON
Canada – 2018



Reiknirit
Er samheiti hvað forrit og kóði þess nákvæmlega gerir

Módel
Er samansafn af kóða eða formúlu eða myndrænu líkani sem getur spá fyrir um niðurstöðu

Neural network
Útreikningar útdelt í hnippi sem tala saman



Í stuttu máli býr gervigreind til
módel sem hún síðan notar til að
spá fyrir um næstu skref!



ChatGPT

(Chat **Generative Pre-trained Transformer**)

Generative

Þýðir að gervigreindin býr til eitthvað. Í þessu tilviki býr til texta.

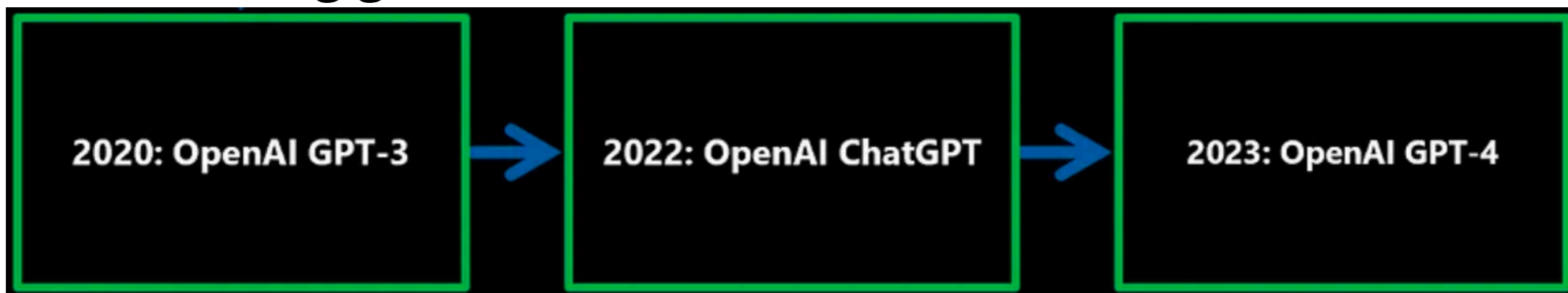
Pre-trained

Byggir á gögnum sem þjálfu upp gervigreindina eða “módelið” t.d. wikipedia til janúar 2022

Transformer

Eitt “lag” af hugsunarferli í neural network módelinu. Þróað af google til að skilja samhengi.

Stöðugt verið að lagfæra og greiða úr villum í módelinu sem liggur að baki notendaviðmótsins





□ **Hluti 2: Rannsóknir læknavísinda**

- Faraldsfræðilegar rannsóknir
- Hraðatakmarkandi skref í rannsóknum
- Hagnýting gervigreindar og takmarkanir



Faraldsfræðilegar rannsóknir

- Notkun á gagnasöfnum yfir marga sjúklinga með skráð 1 (já) eða 0 (nei) um mörg atriði, ásamt kyni, dagsetningum og aldri.
- Gagnasöfn skráð í töflur eða fylki sem þarfnast tölfræðiforrita til að reikna út niðurstöður og varpa á myndrænan hátt.

ID	DOS	Sex	Age	DOA	DOD	LOS	
1	1	2015-04-16	1	99	16.04.2015	17.04.2015	1
2	1	2015-04-16	1	99	16.04.2015	17.04.2015	1
3	2	2012-02-14	2	93	14.02.2012	15.02.2012	1
4	3	2010-02-03	2	88	03.02.2010	06.02.2010	3
5	4	2008-03-11	1	84	11.03.2008	30.05.2008	80
6	5	2015-03-03	2	89	03.03.2015	01.04.2015	29
7	6	2006-06-06	1	78	05.06.2006	15.06.2006	9
8	6	2006-12-05	1	79	28.10.2006	30.03.2007	115
9	7	2007-08-27	2	80	26.08.2007	15.09.2007	19
10	8	2009-12-20	1	82	20.12.2009	06.01.2010	17

```

text=paste("HR:1.08; 95% CI: 0.94 - 1.24; p = 0.3")

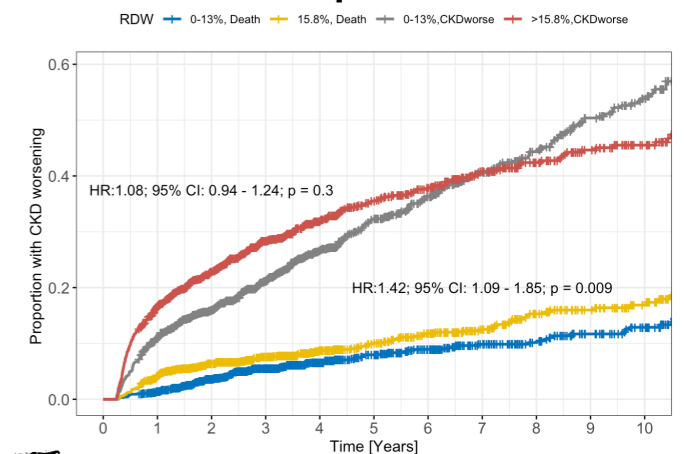
p6$plot=p6$plot+ggplot2::annotate("text",
  x = 2, y = 0.375,
  label = text, size = 4) +theme(axis.text = element_text(size = 12),axis.title.x =
  element_text(color="black", size=12),axis.title.y = element_text(color="black", size=12))

text=paste("HR:1.42; 95% CI: 1.09 - 1.85; p = 0.009")

p6$plot=p6$plot+ggplot2::annotate("text",
  x = 7, y = 0.2,
  label = text, size = 4) +theme(axis.text = element_text(size = 12),axis.title.x =
  element_text(color="black", size=12),axis.title.y = element_text(color="black", size=12))

p6

```



Fylki/Töflur svo umfangsmiklar að ritaður er **kóði** sem framkallar útreikninga fyrir hvern sjúkling í töflunni (84000 talsins) til að fá **niðurstöður**

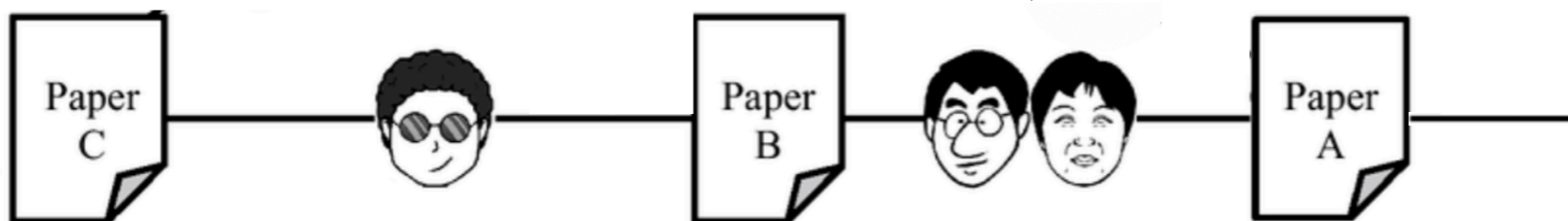
Hraðatakmarkandi skref í rannsóknum



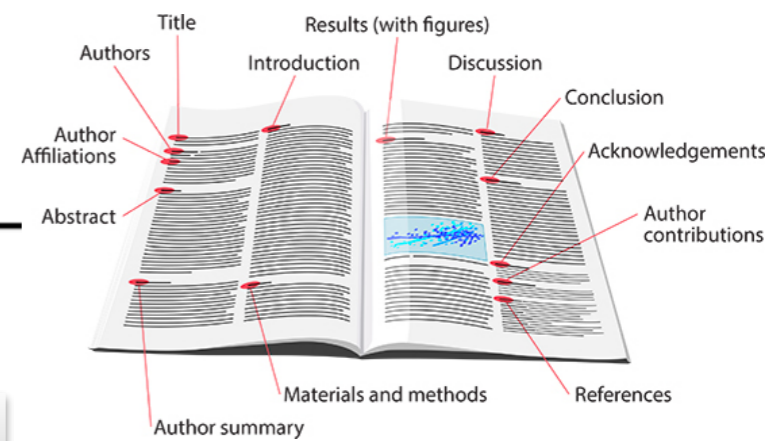
Öflun gagna - tölfræðiútreikninga og úrvinnsla



Ritun vísindagreinar þ.s tvær línur spanna efni heilar blaðsíðu



Yfirferð milli meðhöfunda - innsending til vísindatímarits og leiðréttingar í samræmi





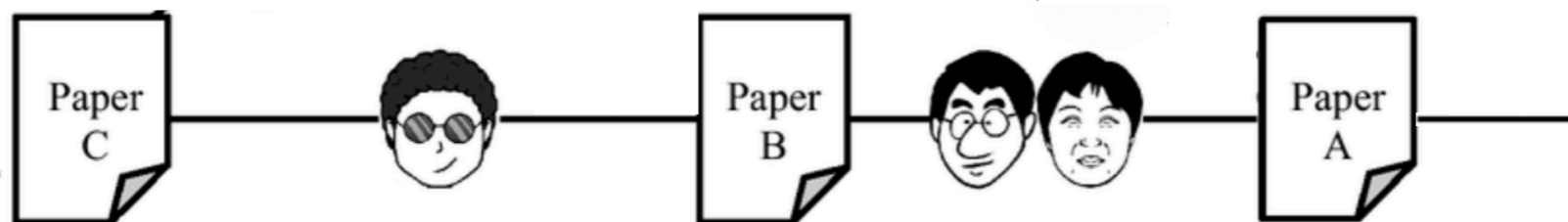
Hraðatakmakandi skref í rannsóknum



Öflun gagna - tölfræðiútreikninga og úrvinnsla

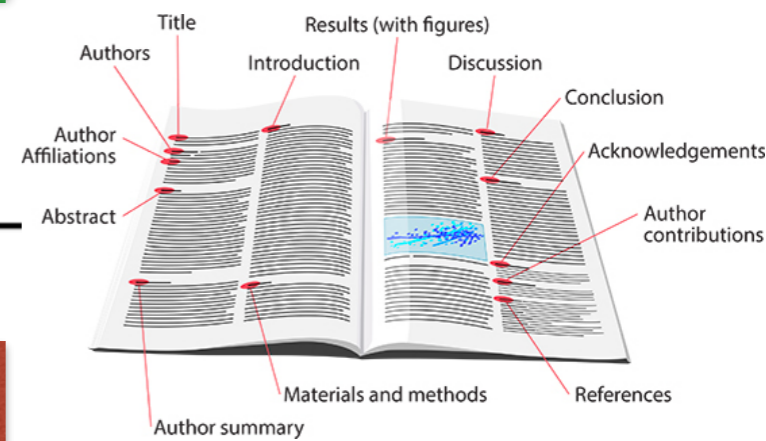


Ritun vísindagreinar þ.s tvær línur spanna efni heilar blaðsíðu



Yfirferð milli meðhöfunda - innsending til vísindatímarits og leiðréttingar í samræmi

Notkun gervigreindar stýttir tíma!





Hagnýting gervigreindar

Tölfræðiútreikningar og úrvinnsla

- Stór gagnasöfn kalla á flókna útreikninga leysta með tölfræðiforritum, þekktast er “R”.
- R-Studio er ritill fyrir R þar sem skrifaður er langur kóði fyrir útreikninga á eigin forritunarmáli.
- Fyrir hverja tegund útreikninga eru sóttir “pakkar” sem hver um sig hefur sitt eigin “orðatiltæki”.

Læra þarf inná kóðunarmál R, og fyrir hvern pakka þarf að læra að nýju “reglur” málsins þar einn vitlaus stafur stöðvar keyrslu kóðans =
TÍMAÞJÓFUR

```
> View(data_matrix)
> data_matrix$colnames
Error in data_matrix$colnames : $ operator
> competing_risks <- cuminc(colnames(da
Error: unexpected '=' in "competing_ris
```



Hagnýting gervigreindar

Tölfræðiútreikningar og úrvinnsla

- ChatGPT hefur hæfileika til að tileinka sér ýmis forritunarmál.
- Hægt er að nota ChatGPT til að læra á R útreikningarpakka og leggja til beinagrind að kóða.

You
make an rstudio code to calculate competing risk from data.frame

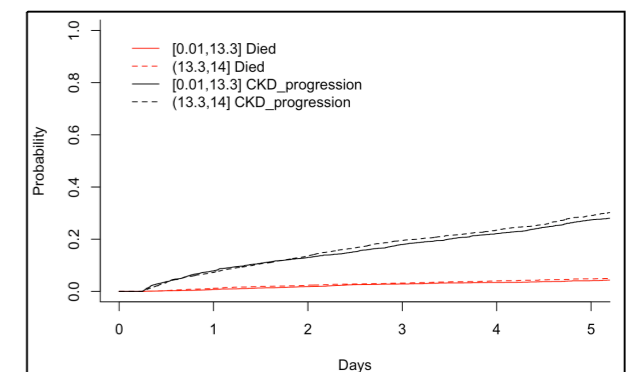
```
# Install and load required packages
if (!requireNamespace("cmprsk", quietly = TRUE)) {
  install.packages("cmprsk")
}
library(cmprsk)

# Create example data frame
data <- data.frame(
  time = c(10, 15, 20, 30, 35, 40, 50, 55),
  status = c(1, 1, 2, 3, 1, 2, 2, 3),
  event = c(1, 2, 2, 1, 2, 1, 2, 1)
)

# Print data frame
print(data)

# Calculate competing risks
competing_risks <- cuminc(data$time, data$status, data$event)
print(competing_risks)
```

```
fit2=cuminc((surgery_sub2$time/365), surgery_sub2$event_f, surgery_sub2$RDWquintiles2, cencode = 0)
print.cuminc(fit2)
```



ChatGPT stytir tíma við kóðagerð en fjarri því að skrifi kóða fyrir hann að öllu leyti!



Hagnýting gervigreindar

Tölfræðiútreikningar og úrvinnsla

You make an rstudio code to calculate competing risk from `matrix`

ChatGPT Sure! Below is an example RStudio code that calculates competing risks using a matrix:

```
R
# Install and load required packages
if (!requireNamespace("cmprsk", quietly = TRUE)) {
  install.packages("cmprsk")
}
library(cmprsk)

# Create example data matrix
data_matrix <- matrix(c(10, 5, 3, 2, 8, 4, 2, 1, 7, 6, 3, 2), nrow = 4, byrow = TRUE)
colnames(data_matrix) <- c("Event A", "Event B", "Event C")
rownames(data_matrix) <- c("Time 1", "Time 2", "Time 3", "Time 4")

# Print data matrix
print(data_matrix)

# Calculate competing risks
competing_risks <- cuminc(data_matrix)
print(competing_risks)
```



You make an rstudio code to calculate competing risk from `data.frame`

ChatGPT Certainly! Below is an example RStudio code to calculate competing risks using a data frame:

```
R
# Install and load required packages
if (!requireNamespace("cmprsk", quietly = TRUE)) {
  install.packages("cmprsk")
}
library(cmprsk)

# Create example data frame
data <- data.frame(
  time = c(10, 15, 20, 30, 35, 40, 50, 55),
  status = c(1, 1, 2, 3, 1, 2, 2, 3),
  event = c(1, 2, 2, 1, 2, 1, 2, 1)
)

# Print data frame
print(data)

# Calculate competing risks
```





Hagnýting gervigreindar

Tölfræðiútreikningar og úrvinnsla

ChatGPT gerir flókna kóða og gerir oft illa greinarmun á einföldum atriðum sem skipta máli í ritun kóða - því er grundvallarskilningur á kóðunarmálinu áfram nauðsynlegur



Ritun vísindagreinar þ.s tvær línur spanna efni heilar blaðsíðu

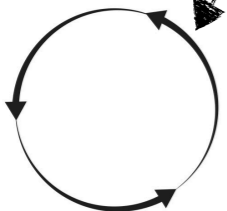
Hagnýting gervigreindar

Ritun vísindagreinar

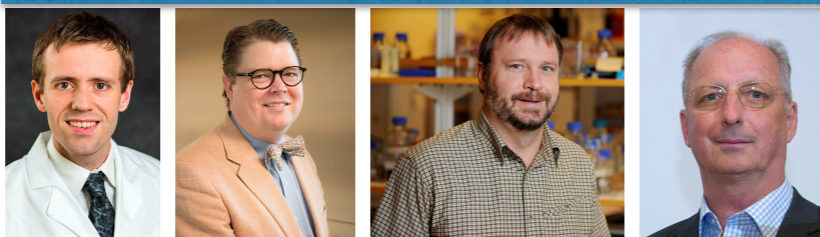
- Orðafjöldi takmarkaður af vísindatímariti við ritun greina. Fá orð þurfa að skýra margt.
- Rík krafa um að “stytta” texta umfram það sem talið gerlegt í upphafi - en texti innihaldi sömu upplýsingar.

One possible explanation of relationship between elevated RDW and increased risk of mortality is the underlying response of blood tissue to inflammation. Inflammatory cytokines inhibit the development of red blood cells in the marrow causing imminent anemia²⁴, the main compensation for this has been shown to be extension the lifetime of red blood cells²⁵. Hence the patients with elevated RDW pre-operatively may already be compensating for potential underlying illness, already stretching the capability of the blood tissue to sustain normal function. Supporting this theory is the finding of

81 orð



4 Paulreyndir vísindamenn



Martin Ingi Sigurðsson Prófessor og yfirlæknis í svæfinga- og gjörgæslulækningum við læknadeild Háskóla Íslands
Kenneth B. Christopher Aðstoðarprófessor við Harvard Medical School og nyrnalæknir
Sigurbergur Kárason Yfirlæknir svæfinga- og gjörgæslulækninga á landspítala og rannsakandi.
Gísli H. Sigurðsson Prófessor í svæfinga- og gjörgæslulækningum



matching, and there was an evidence of dose response and a consistency of the magnitude of the relationship in multiple subgroups of patients. This suggests that in surgical cohorts increased RDW might indicate a further burden of chronic illness not captured by preexisting known comorbidities, and this signal warrants further investigation.

22 orð

Ritun vísindagreinar þ.s tvær línur spanna efni heilar blaðsíðu

Hagnýting gervigreindar

Ritun vísindagreinar

- Notkun gervigreindar til að stytta flókna texta kemst nálægt sömu afköst og 4 vísindamenn
- Yfirleitt kemur með tillögu að styttingu sem höfundur og meðhöfundar breyta síðar
- Sé notað rétt getur stytt tímabátt við ritun greina umtalsvert.

One possible explanation of relationship between elevated RDW and increased risk of mortality is the underlying response of blood tissue to inflammation. Inflammatory cytokines inhibit the development of red blood cells in the marrow causing imminent anemia²⁴, the main compensation for this has been shown to be extension the lifetime of red blood cells²⁵. Hence the patients with elevated RDW pre-operatively may already be compensating for potential underlying illness, already stretching the capability of the blood tissue to sustain normal function. Supporting this theory is the finding of

81 orð



ChatGPT
"Elevated RDW may reflect inflammation-driven blood tissue response, potentially from underlying disease, leading to anemia and compensatory red blood cell extension."

22 orð

matching, and there was an evidence of dose response and a consistency of the magnitude of the relationship in multiple subgroups of patients. This suggests that in surgical cohorts increased RDW might indicate a further burden of chronic illness not captured by preexisting known comorbidities, and this signal warrants further investigation.



Martin Ingi Sigurðsson Prófessor og yfirlæknis í svæfinga- og gjörgæslulækningum við læknadeild Háskóla Íslands
Kenneth B. Christopher Aðstoðarprófessor við Harvard Medical School og nýrnalæknir
Sigurbergur Kárason Yfirlæknir svæfinga- og gjörgæslulækninga á landspítala og rannsakandi.
Gísli H. Sigurðsson Prófessor í svæfinga- og gjörgæslulækningum

Hagnýting gervigreindar

Gervigreind nýtanlegt tól, ekki höfundur

Two US lawyers fined for submitting fake court citations from ChatGPT

Law firm also penalised after chatbot invented six legal cases that were then used in an aviation injury claim

Nearly 400 uni students investigated for using ChatGPT to plagiarise assignments

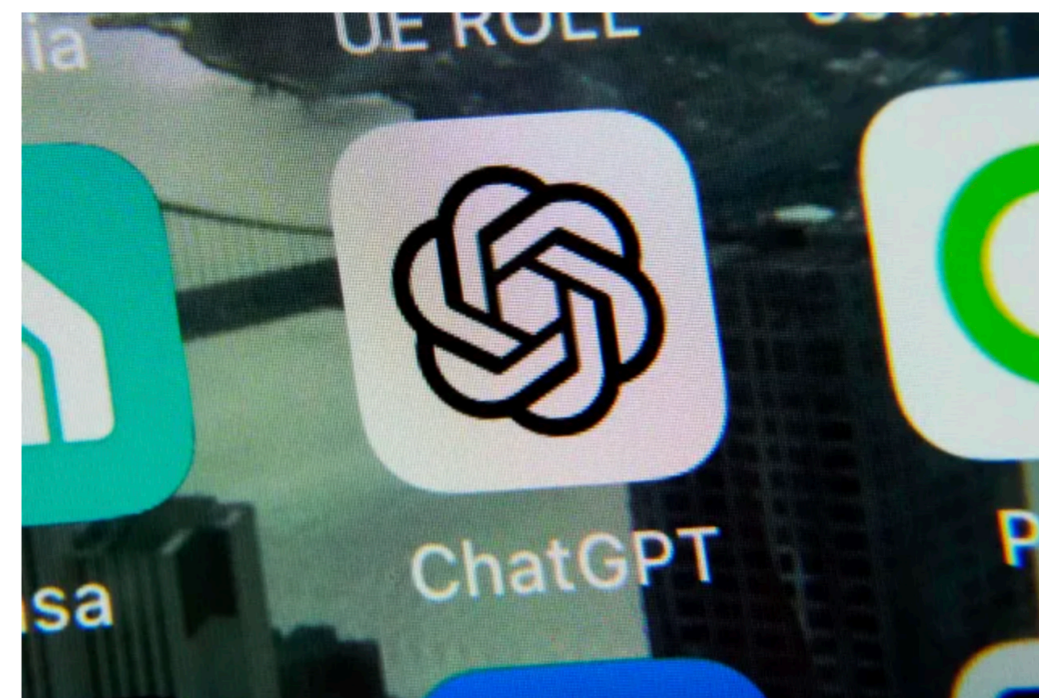
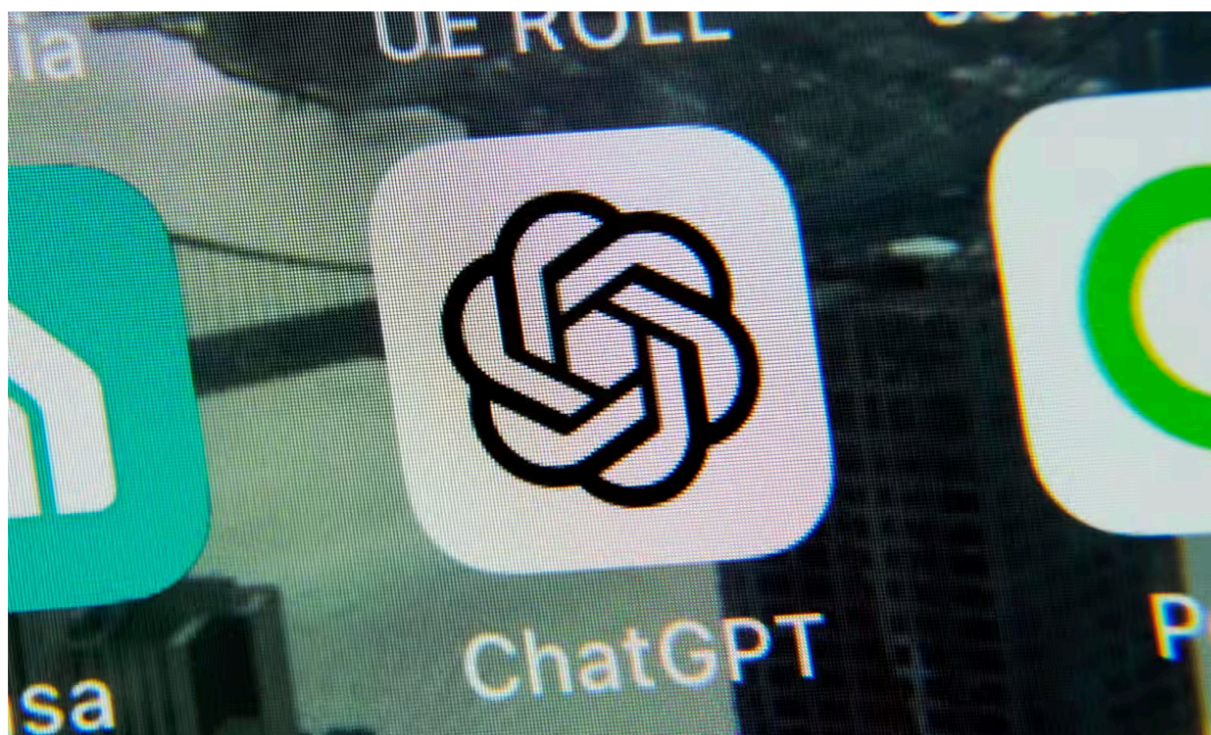


Tom Sanders

Published Jul 5, 2023, 6:11pm | Updated Jul 5, 2023, 6:18pm



Comment





Samantekt

- Gervigreind er í stuttu máli módel sem tekur gögn og spáir fyrir um niðurstöðu (málfræðileg svör, tala)
- ChatGPT er málfræðimódel og getur tileinkað sér kóðunarmál tölfræðiforrita.
- Faraldsfræðilegar rannsóknir svo gagnamiklar að flókinn kóði í tölfræðiforritum þarf til úrvinnslu
- ChatGPT gagnast við að rita beinagrind að tölfræðikóða og er sérstaklega gott í að stytta flóknar málsgreinar, síður að starfa sem rithöfundur.

Takk fyrir mig!