Priručnik za preživljavanje?

- Megatrend poslovna rješenja već 30 godina radi u IT industriji
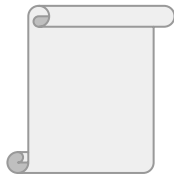
1. O metodologiji i drugim nužnim zlima
Ili: O nekim pogreškama koje sam učinio

2. Kauzalna analiza
Kako polako počinjemo formalizirati uzroke i posljedice

3. Demo
Napokon nešto što se može kopipejstati!

O metodologiji i drugim nužnim zlima

> *"All statistical modeling has these same two frames: the small world of the model itself and the large world we hope to deploy the model in. Navigating between these two worlds remains a central challenge of statistical modeling. The challenge is aggravated by forgetting the distinction."*
>
> *Richard McElreath*

- "Our search identified 31 587 studies, of which 82 [...] were included."
- "69 studies provided enough data to construct contingency tables, enabling calculation of test accuracy…"
- "An out-of-sample external validation was done in 25 studies"
- "… of which 14 made the comparison between deep learning models and health-care professionals in the same sample"

- "Our search identified ▦ studies"
- "of which ▦ were included after initial screening"
- "after quality screening, ▦ studies were included in this systematic review"
- "Our review finds that ▦ of the models identified are of potential clinical use due to methodological flaws and/or underlying biases."

- "Our search identified 2,212 studies"
- "of which 415 were included after initial screening"
- "after quality screening, 62 studies were included in this systematic review"
- "Our review finds that none of the models identified are of potential clinical use due to methodological flaws and/or underlying biases."

Nedostatna dokumentacija o:

- ◻ Tehnikama za redukciju dimenzionalnosti (52%)

- ◻ Validaciji modela (61%)

Nedostatna dokumentacija o:

- ▫ Tome kako je konačan model izabran (61%)

- ▫ Metodama pretprocesiranja slika (58%)

- ▫ Detalji trening procedure (npr. Optimizator, funkcija gubitka, hiperparametri) (49%)

Detalji o modelu su bitni!

Za replikaciju su neophodni detalji o proceduri selekcije modela, metode pretprocesiranja, hiperparametri, itd.

Od 37 uvrštenih:

- ▫ ■ ih nije napravilo bilo kakvu eksternu validaciju

- ▫ ■ ih nije koristilo statističke testove kako bi procijenili značajnost ili odredili confidence intervale

- ▫ ■ ih nije dalo informacije o demografiji u različitim skupovima podataka

Od 37 uvrštenih:

- 29 ih nije napravilo bilo kakvu eksternu validaciju

- 26 ih nije koristilo statističke testove kako bi procijenili značajnost ili odredili confidence intervale

- 26 ih nije dalo informacije o demografiji u različitim skupovima podataka

# Eksterna validacija je neophodna.

Ima jedna priča o tenkovima i noći...

- Nije mu podložna samo umjetna inteligencija: anti-tenk psi!

- Dvije bolnice specijaliziraju se za različite bolesti i svaka ima specifičnu anotaciju na npr. rendgenu.

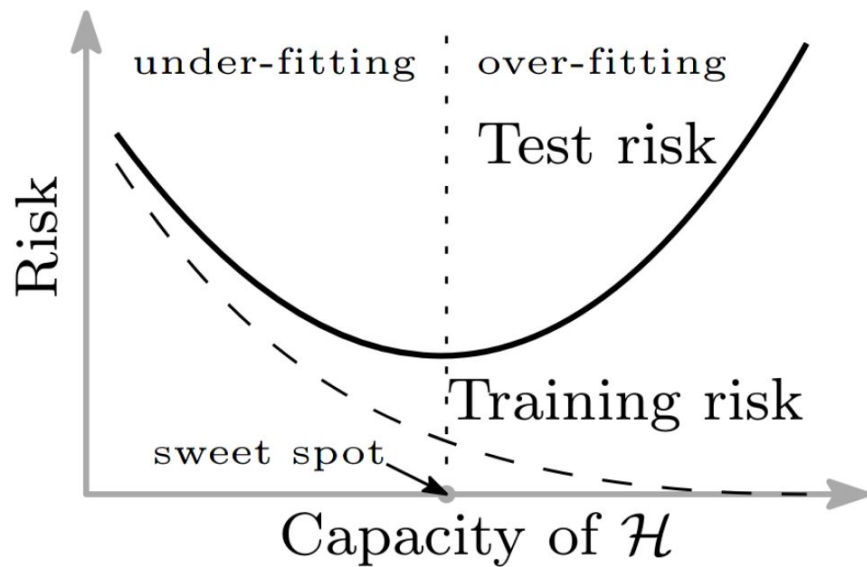- CNN za problematične lezije na koži: ravnalo kao feature.

- "Ten papers used cross-validation to evaluate model performance…"

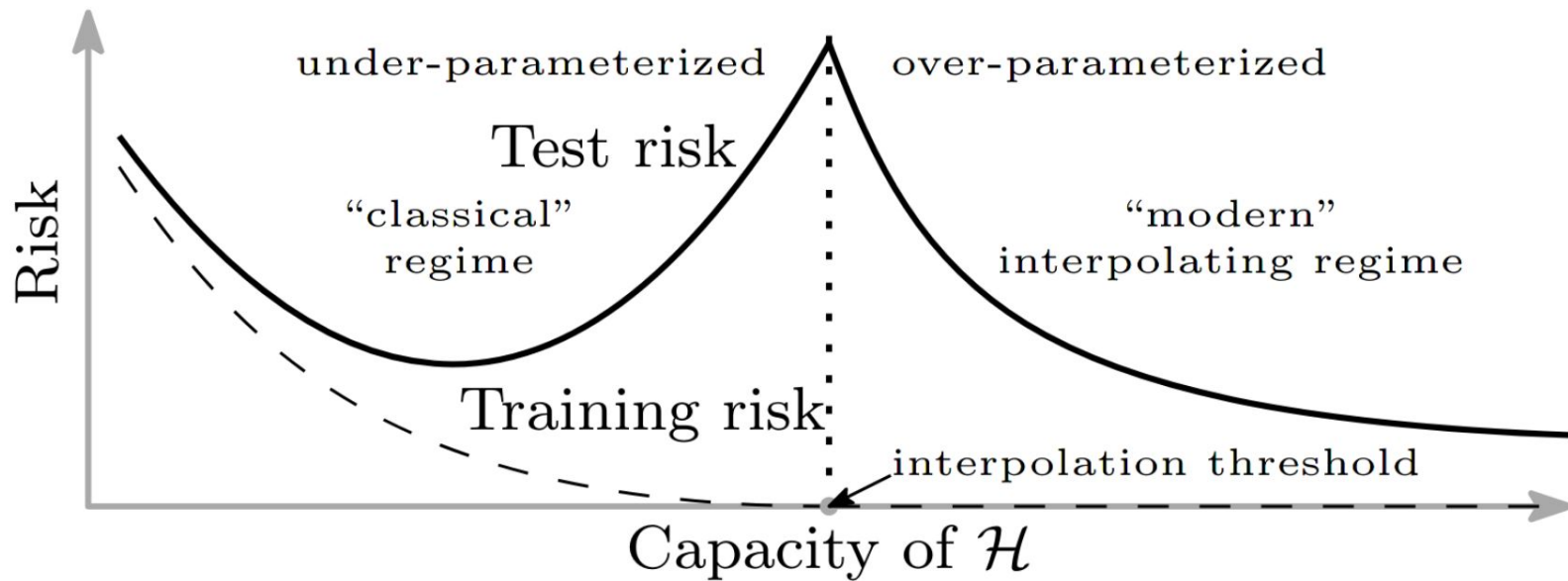- "… one used both cross-validation and an external test set"

- Itd. :(
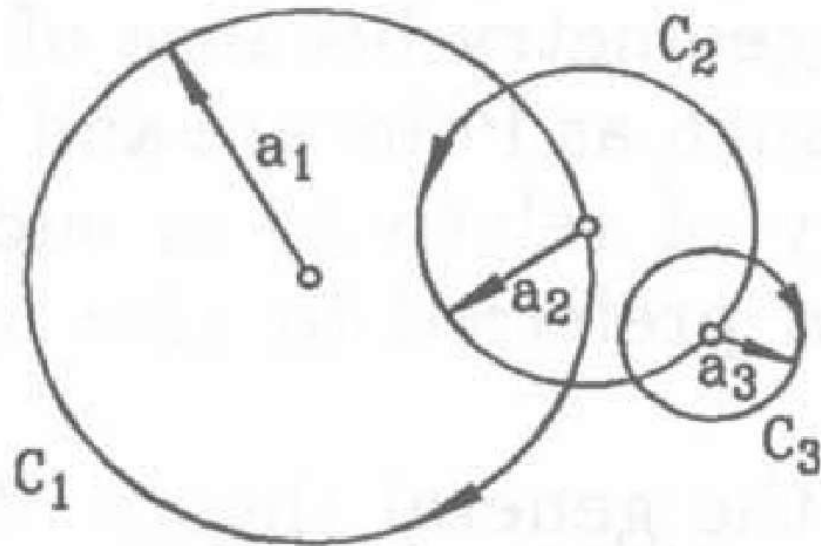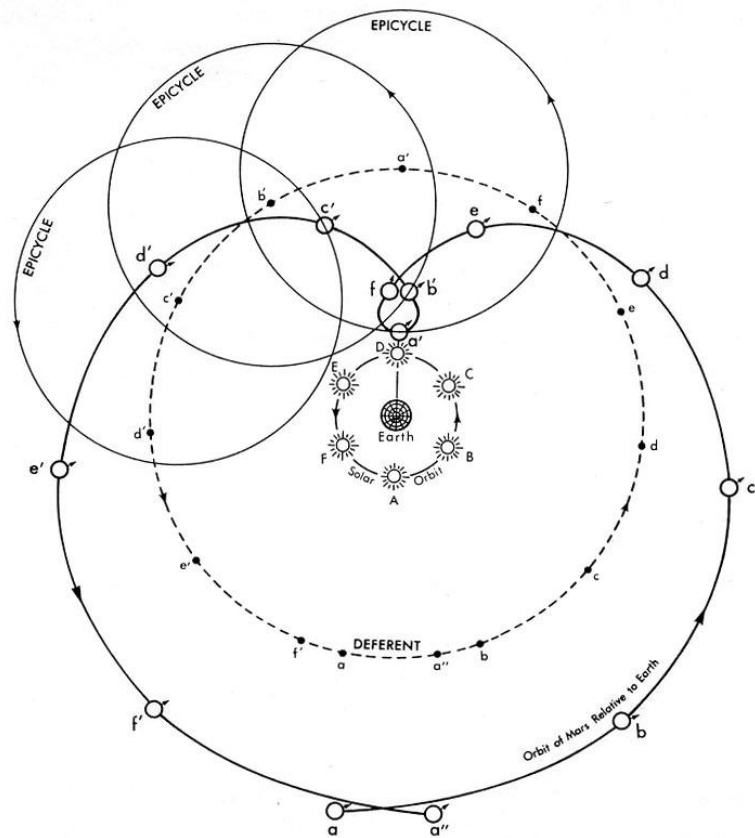
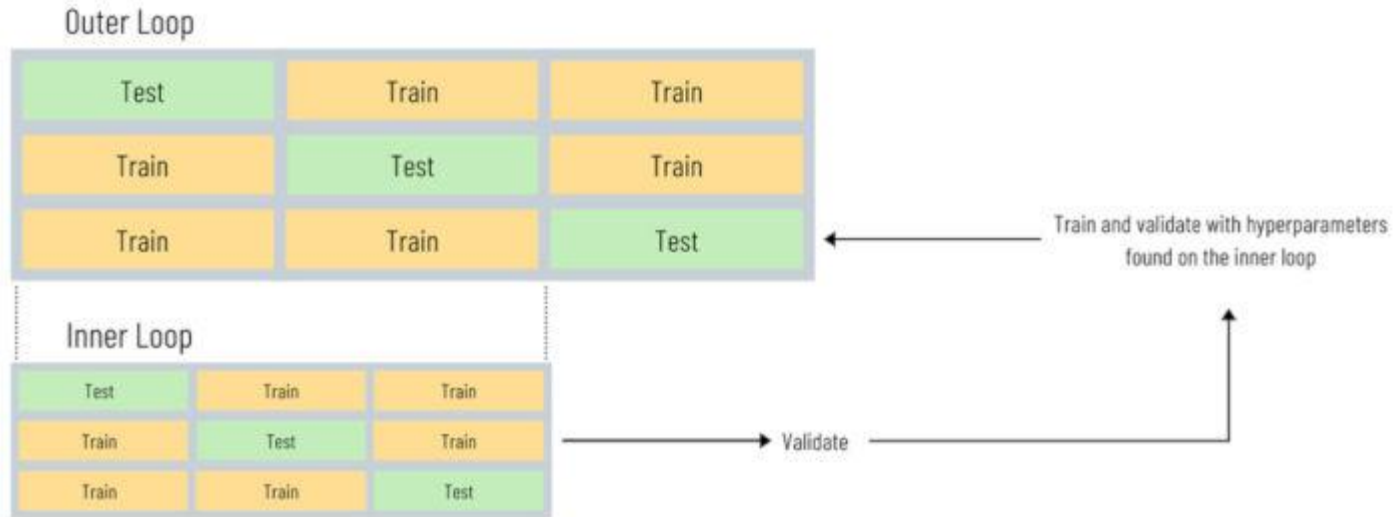Idealan workflow:

- Jednostavniji-nego-što-zamišljate baseline

- Malo-kompleksniji-ali-još-jednostavniji-nego-što-zamišljate baseline

- Linearna/logistička regresija – razumjeti zašto failaju

- Kompleksniji modeli

Repeat for **n** rounds

Outer Loop

Test	Train	Train
Train	Test	Train
Train	Train	Test

← Train and validate with hyperparameters found on the inner loop

Inner Loop

Test	Train	Train
Train	Test	Train
Train	Train	Test

→ Validate

- Reproducibility project na Sveučilištu u Virginiji je, od 2011. do 2015., pokušao replicirati 100 studija u psihologiji. [Samo ili čak?] ███ ih se repliciralo.
- Po određenim statistikama oko ███% machine learning projekata nikada ne dođe u produkciju.

- Reproducibility project na Sveučilištu u Virginiji je, od 2011. do 2015., pokušao replicirati 100 studija u psihologiji. [Samo ili čak?] 35 ih se repliciralo.
- Po određenim statistikama oko 90% machine learning projekata nikada ne dođe u produkciju.

Neurons in the mouse brain correlate with cryptocurrency price: a cautionary tale

Guido T. Meijer

Champalimaud Centre for the Unkown, Lisbon, Portugal
Correspondence: guido.meijer@research.fchampalimaud.org

24 April 2021

Abstract

In this paper I report the discovery of neurons which showed a neural correlate with ongoing fluctuations of Bitcoin and Ethereum prices at the time of the recording. I used the publicly available dataset of Neuropixel recordings by the Allen Institute to correlate the firing rate of single neurons with cryptocurrency price. Out of $\sim$40.000 recorded single neurons, $\sim$70% showed a significant correlation with Bitcoin or Ethereum prices. Even when using the conservative Bonferroni correction for multiple comparisons, $\sim$35% of neurons showed a significant correlation, which is well above the expected false positive rate of 5%. These results were due to 'nonsense correlations': when correlating two signals which both evolve slowly over time, the chances of finding a significant correlation between the two are much higher than when comparing signals which lack this property.

"

"The reviewer, the funder, and the person on the street, all think that defensiveness is how you do science; and how you manage, and how you make law - that the essence of civilization is to avoid things being criticizable."

Eliezer Yudkowsky

2

Kauzalna analiza

"If I could sum up the message of this book in one pithy phrase, it would be that you are smarter than your data. Data do not understand causes and effects; humans do."

Judea Pearl

- Uspoređujemo dva tretmana (A i B) za liječenje bubrežnih kamenaca.

- Tretman A bolje liječi (ima veću stopu uspješnosti) za veće kamence.
- Tretman A bolje liječi (ima veću stopu uspješnosti) za manje kamence.

- Tretman B bolje liječi (ima veću stopu uspješnosti) kada se gledaju i manji i veći kamenci.

	Tretman A	Tretman B
Mali kamenac	**81 od 87 (93%)**	**234 od 270 (87%)**
Veliki kamenac	**192 od 263 (73%)**	**55 od 80 (69%)**
Ukupno	**273 od 350 (78%)**	**289 od 350 (83%)**

	Tretman A	Tretman B
Mali kamenac	**99 od 100 (99%)**	**98999 od 100000 (~99%)**
Veliki kamenac	**1000 od 100000 (1%)**	**1 od 101 (~1%)**
Ukupno	**1099 od 100100 (~1%)**	**99000 od 100101 (~99%)**

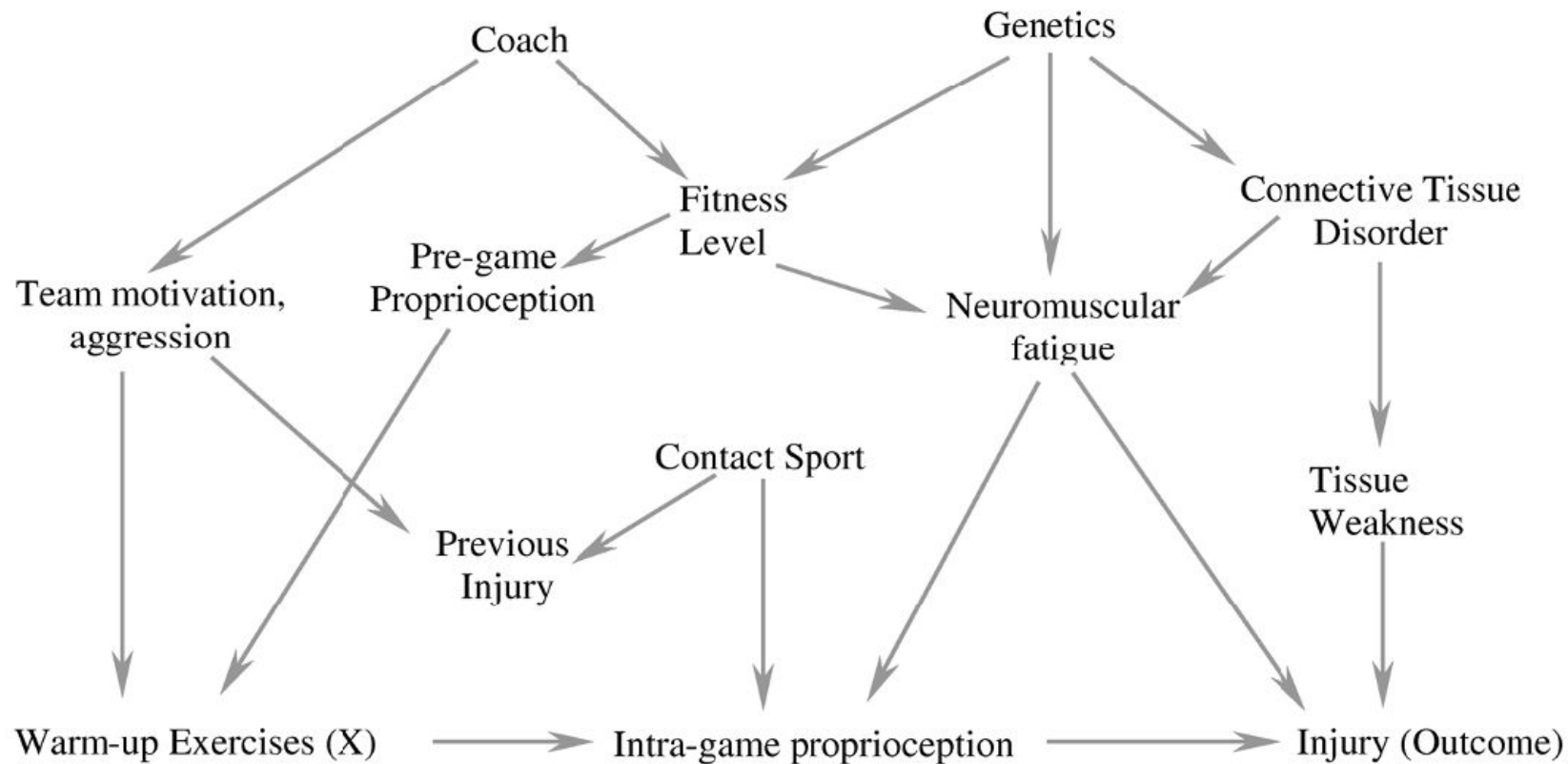# Zašto kauzalna analiza?

Jer klasična statistika / ML nemaju mehanizme za odgovaranja na pitanja "zašto?" i "što ako?".
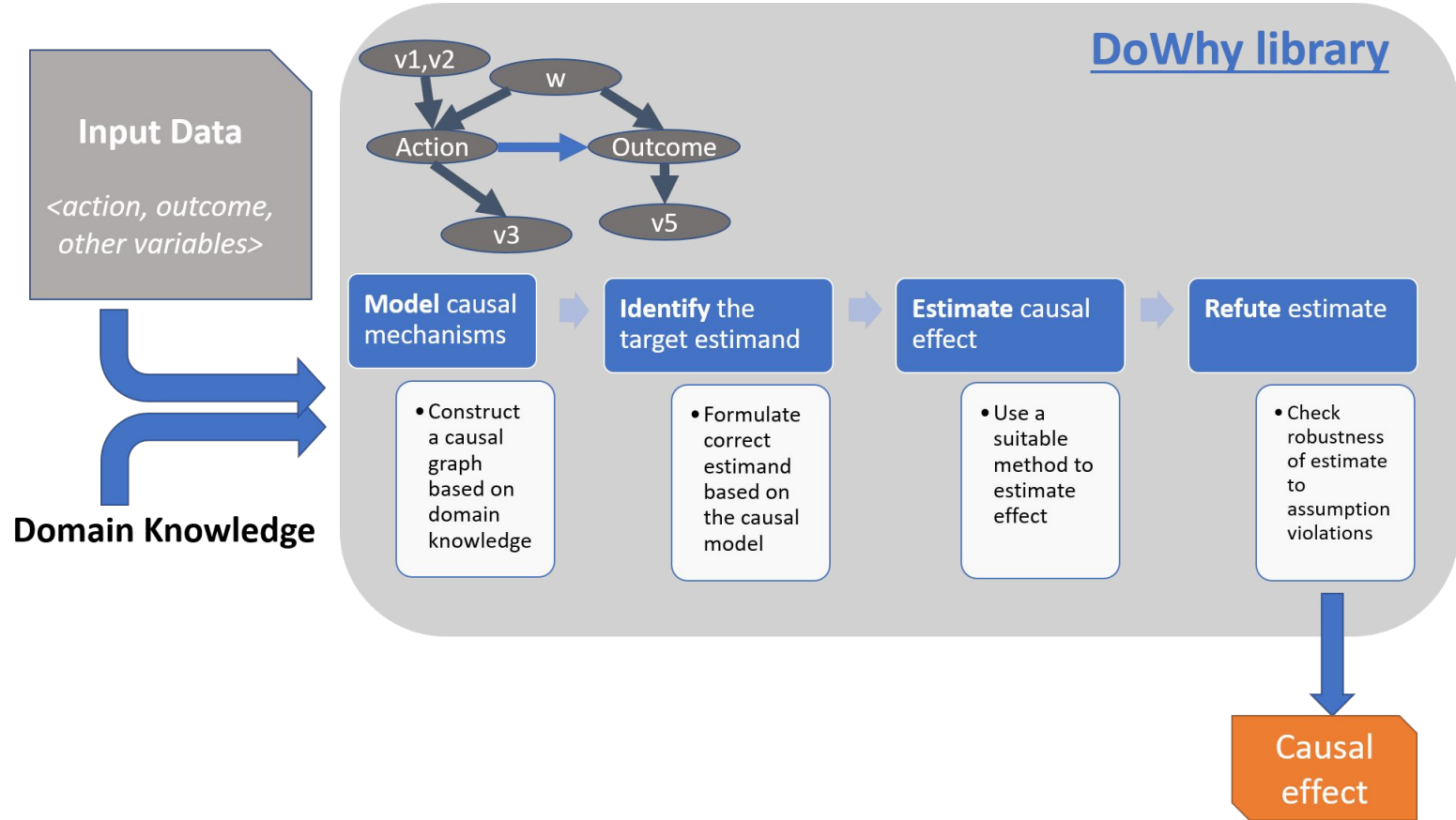
A

↓

B

↓

C

1. A i B su zavisni
   - Za neke a, b:
   - $P(B = b | A=a) \mathrel{!=} P(B=b)$
2. B i C su zavisni
   - Za neke b, c:
   - $P(C=c | B=b) \mathrel{!=} P(C=c)$
3. A i C su vjerojatno zavisni
   - Za neke a, c:
   - $P(C=c \mid A=a) \mathrel{!=} P(C=c)$
4. C i A su nezavisni, uvjetno na B
   - Za sve a, b, c:
   - $P(C=c | B=b, A=a) = P(C=c | B=b)$

3

Demo!

Repozitorij u kojem je kod, prezentacija, reference (koje pusham uskoro), je na:

- https://github.com/rovle

Pitanja, komentari, želje? :)

Ako ste zainteresirani za stručnu praksu, javite se na
lovre.pesut@megatrend.com

(Ako ste zainteresirani za spašavanje svijeta, javite se na
lovre@efektivnialtruizam.org)