

# Втор парцијален испит по предметот Методологија на истражување во ИКТ

Давид Галевски – 171194

За првиот колоквиум правев тетратка за EEG feature extraction за детекција на P300 сигнали за детекција на лаги (Abootalebi, Moradi, Khalilzadeh, 2009).

## 2. Објаснете ја разликата помеѓу репродусибилност и репликабилност

Овие 2 термини се слични еден со друг, и во литературата често се користат наизменично. Во главно и 2та концепти се обидуваат да ги одговорат следниве прашања (National Academies of Sciences, Engineering, Medicine, 2009):

- Дали податоците и методите за анализа на податоците се јавно достапни (достапен код и податоци), за да може резултатите (од трудот) лесно да се проверат од независни лица.
- Доколку анализата се изврши одново дали се добиваат истите резултати, користејќи ги истите податоци?
- Дали може да се донесат истите заклучоци користејќи нови податоци?

Во компјутерски науки, репродусибилно истражување ги одговара првите две прашања. Кодот и податоците од истражувањето се јавно достапни (најчесто на github), се користи docker, има requirements.txt фајл итн...

Репликабилноста се однесува на третото прашање. Ако се соберат нови податоци и се обидеме да го одговориме истото прашање од трудот, дали ќе ги добиеме истите резултати? Доколку одговорот на ова е да, нашиот труд е репликабилен.

Околу 2010та година се појавува нова криза наречена „replication crisis“, поради тоа што најголем дел од објавените трудови (најчесто од областа на social sciences и медицина) не се репликабилни. Ова е голема криза бидејќи репликабилноста е клучен дел од научниот метод.

### 3. Која е разликата помеѓу контејнер и виртуелна машина?

Виртуелни машини се токму тоа, виртуелни компјутери со виртуелен хардвер на кој што има оперативен систем различен од оној на хост машината. Виртуелните машини се доста големи бидејќи за секоја виртуелна машина има посебни библиотеки, апликации и фајлови, па тие може да зафаќаат неколку гигабајти.

Токму тоа што виртуелните машини имаат посебни оперативни системи во изолирана околина додава премногу overhead и троши многу ресурси.

Контејнерите се доста помали и полесни од виртуелните машини. Тие работат над хост оперативниот систем, го користат неговиот кернел, библиотеки и binaries. Поради тоа тие не мора да го репродуцираат цел код од оперативниот систем па се значително полесни, и се многу лесни за одржување.

Целта за која што се користат и контејнери и виртуелни машини е слична, и 2та најчесто се користат за да понудат идентична околина за правење на одредени пресметки без разлика од околината на хост компјутерот.

## References

Abootalebi V, Moradi M.H, Khalilzadeh M.A, (2009) A new approach for EEG feature extraction in P300-based lie detection. Computer Methods and Programs in Biomedicine,

<https://doi.org/10.1016/j.cmpb.2008.10.001>

National Academies of Sciences, Engineering, and Medicine. (2019). Reproducibility and Replicability in Science. Washington (DC): National Academies Press (US)

<https://www.ncbi.nlm.nih.gov/books/NBK547546/>