Matematički fakultet

Univerzitet u Beogradu

Seminarski rad:

**Uticaj P-adičnosti na razilke genetskog koda SARS1, SARS2 i MERS koronavirusa**

Mentor: Studenti:

Prof. dr Nenad Mitić Marko Koprivica Milan Bodo

[**1.Uvod 3**](#_tlswcz3ba9a2)

[**2.P-adična distanca 3**](#_25pa96rcljka)

[**2.1 P-adično kodiranje nukleotidnih sekvenci 3**](#_vrt2f53qlwmw)

[2.2 P-adična distanca izmedju kodona 3](#_cheduihhfouv)

[**3.Podaci 4**](#_kq2dav6odgbr)

[**4. Modeli 4**](#_i2ebtys0g6mv)

[**5. Zaključak 4**](#_nsqyetom2fvg)

[**6.Reference 5**](#_hhbjuij9zxr0)

# 1.Uvod

Ovaj seminarski rad bavi se uticajem p-adičnosti na različitosti genetskog koda kod SARS1, MERS i SARS2 virusa, i to na primerima klasifikacije ovih virusa, klasifikacije proteina virusa i klasifikacije sojeva SARS2 virusa po WHO klasifikaciji.

# 2.P-adična distanca

Kako se ∀m∈Z , m≠0 može zapisati kao m=pk•а gde je p prost broj, k ∈{0,1,2,...} i a ceo broj koji nije delji sa p. P-adična distanca broja m u oznaci |m|p  se definiše na sledeći način:

|m|p ≔p-k [1]

i specijalno za m=0, |m|p≔0.

P-adičcna distanca dva cela broja x i y se racuna na sledeći način:

|x-y|p

P-adičcna distanca je ultrametrika, to jest za nju važi:

1. |x-y|p ≥ 0 , |x-y|p＝0 ⇔x＝y

2. |x-y|p＝|y-x|p

3. |x-y|p ≤ max {|x-z|p ,|y-z|p}

Primer:

|63-3|2= |60|2= |22•3•5|2= ½

|63-3|3= |60|3= |22•3•5|3= ⅓

## 2.1 P-adično kodiranje nukleotidnih sekvenci

Nukleotidna sekvenca se deli u kodone(niske od po tri nukleotida), koji se kodiraju u trocifrene brojeve, gde se svakom od nukleotida ({A,C,G i T}) pridružuje po jedan broj iz Z5={0,1,2,3,4}, i to na sledeći način:

C=1 A=2 U=3 G=4

Primetimo da se purinske aminokiseline(A,G) i pirinske aminokisleine(C,U) kodiraju kao cifre iste parnosti, što ćemo korisititi u računanju distanci izmedju kodona.

## 2.2 P-adična distanca izmedju kodona

# 3.Podaci

Istraživanje smo izvodili na bazi podataka od 13202 instanci. Najvažniji atributi su ime virusa, ime proteina, i nukleotidna sekvenca.

| Virus | Broj instanci |
| --- | --- |
| SARS1 | 18 |
| MERS | 1859 |
| SARS2 | 11325 |

Tabela 3.1 raspodela instanci po tipu virusa

Još za SARS2 raazlikovali su se i sojevi virusa:

| Soj SARS2 virusa | Broj instanci |
| --- | --- |
| Alpha, Delta, Epsilon, Gamma, Iota, Omricon | 1500 |
| Eta | 628 |
| Beta | 586 |
| Zeta | 467 |
| Lambda | 410 |
| Kappa | 157 |
| Mu | 65 |
| Theta | 20 |

Tabela 3.1 raspodela instanci po soju SARS2 virusa

Sve sekvence proteina su jedinstvene, nukleotidno kompletne i bez višeznačnih karaktera.

# 4. Modeli

# 5. Zaključak

# 6.Reference

Dragovich,, Branko, and Nataša Ž Mišić. “P-Adic Hierarchical Properties of the Genetic Code.” *BioSystems,*, vol. 185, no. 104017, 2019, [1]