

Random Forest modellek a rákkutatásban

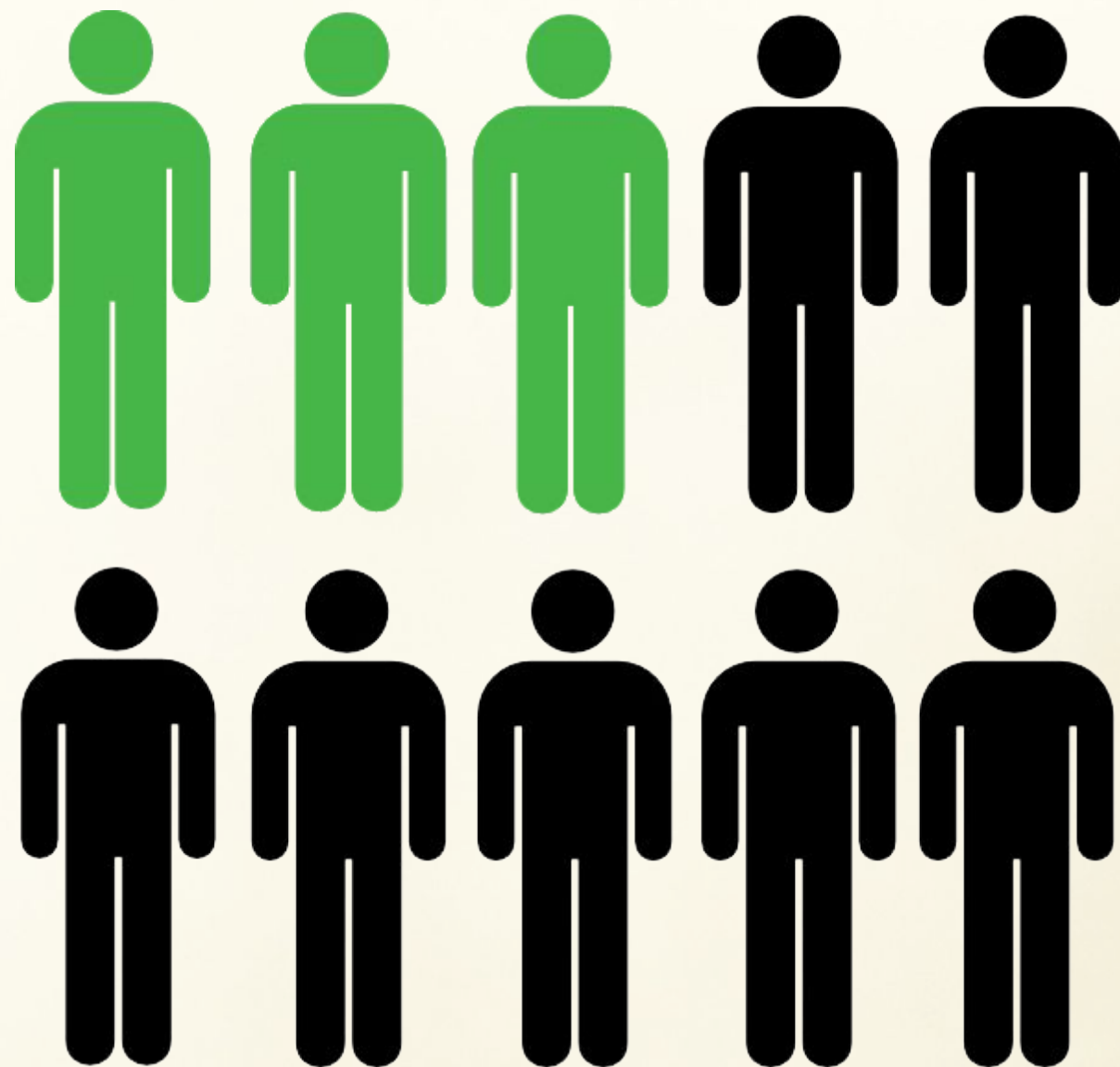
Papp Benjamin Tamás













Gyulladt / kötőszövetes



Gyulladt

IMMUN FENOTÍPUS

Sivatag



Bagaev et.al. 2021 Cancer Cell

Gyulladt / kötőszövetes



IMMUN FENOTÍPUS



Sivatag



Bagaev et.al. 2021 Cancer Cell



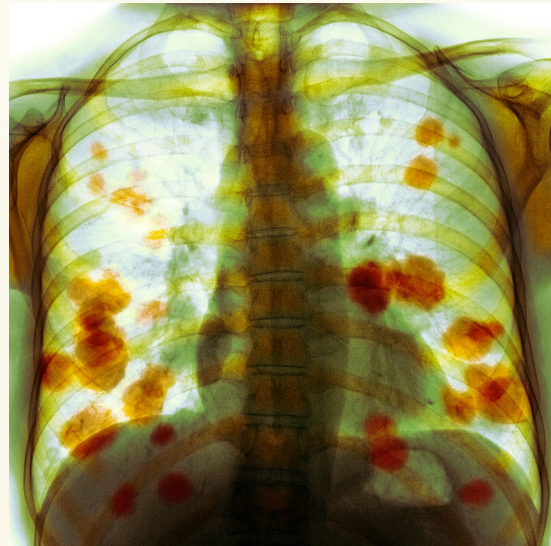
~10 000 tumoros minta

435 melanomás minta

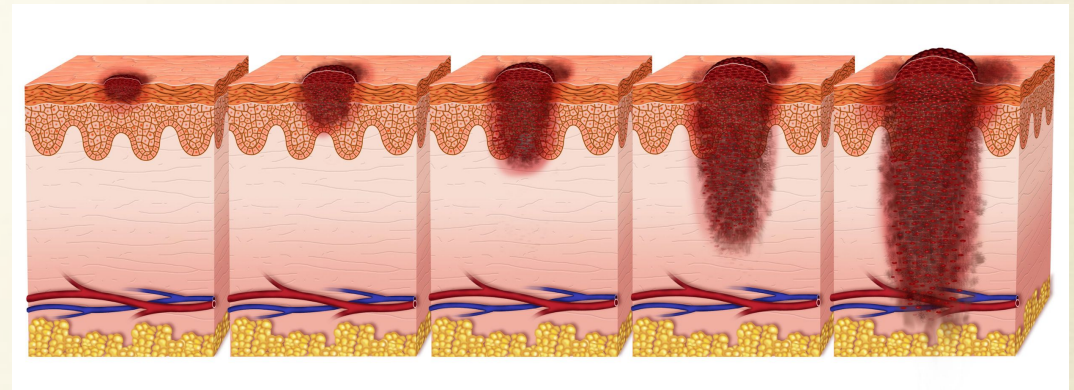
Demográfia

- Kor
- Nem
- BMI

Áttét?



Mélybe terjedés (T)





~10 000 tumoros beteg

→ 435 melanomás minta

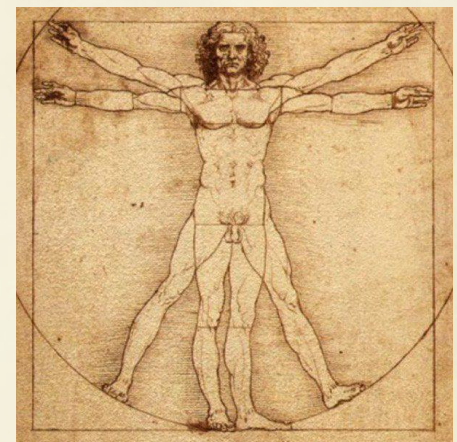
Nyirokcsomóáttét



Visszatérő?



Lokalizáció



Prediktor változók

- Folytonos változók: Kor, BMI
- Bináris változók: Nem, Áttét-e?, visszatérő?
- Kategorikus változók □ dummy variable: Mélybe terjedés, nyirokcsomóba terjedés, lokalizáció
- Összesen 21 prediktor változó

Kimeneti változók

- GY (n=118)
- GYK (n=63)
- K (n=83)
- S(n=171)



GY vs. GYK	GYK vs. K
GY vs. K	GYK vs. S
GY vs. S	K vs. S

Első modellek

- Irodalmi példa alapján: Tanítási –
teszt szett 80%-20% arányban
- Teszt szett prediktálása □ ROC AUC
értékek
- 10 000 iteráció
- Medián, max ROC AUC értékek

	ROC AUC értékek, minden változó	
	Median	Max
GY vs. GYK	0.612	0.913
GY vs. K	0.664	0.914
GY vs. S	0.667	0.851
GYK vs. K	0.619	0.941
GYK vs. S	0.602	0.869
K vs. S	0.596	0.857

Dimenziócsökkentés

- VSURF algoritmus
- Minden modell „típusra” kölön futtatva

	ROC AUC értékek, minden változó		ROC AUC értékek, VSURF változók	
	Median	Max	Median	Max
GY vs. GYK	0.612	0.913	0.711	0.98
GY vs. K	0.664	0.914	0.703	0.946
GY vs. S	0.667	0.851	0.73	0.923
GYK vs. K	0.619	0.941	0.656	0.918
GYK vs. S	0.602	0.869	0.653	0.902
K vs. S	0.596	0.857	0.621	0.884

[illegible]

WHAT
TO DO





ImmunoInfo

Systems Immunology Research Group



papp.benjamin.t@gmail.com