



STARSCHEMA

Hogyan kerüljünk a legjobb tíz közé egy Kaggle versenyen?

Budapest ML Fórum

- Tóth Benedek
- Zempléni Balázs



Agenda

- A verseny bemutatás
- Előfeldolgozás, addicionális információk hasznosítása
- A neurális hálók részletei
- Különböző modellek összefésülése
- Más csapatok megoldásai
- Konklúziók

Csapattagok

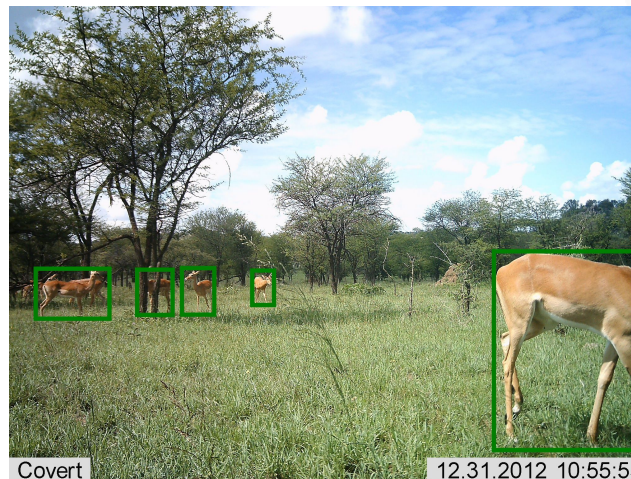
- Barthó Gergely
- Fekete Ákos
- Kapronczay Mór
- Tóth Benedek
- Windhager-Pokol Eszter
- Zempléni Balázs

Röviden a versenyről

- **Input:** vadonban kihelyezett “kameracsapdák” képsorozatai
- **Cél:** a képeken szereplő állatfajok beazonosítása
- **Kihívás:** új helyszínekre is generalizálható modell építése

Részletek

- 217 959 tanító kép 441 helyszínről
- 62 894 tesztkép 111 helyszínről
- 39-406 műholdkép per helyszín
- Előre tanított “állatfelismerő” modell: MegaDetector (Faster-RCNN model with Inception-Resnet-v2 backbone)



Képek előfeldolgozása, data augmentation



90° Rotate



Crop



Rotation



Brightness



Exposure



Blur



Noise

Hisztogram-kiegyenlítés

- A kép kontrasztjának megnövelése a színek komponensek eloszlásának uniformizálásával
- Adaptív változat: minden pixel a saját környezetére alapján kerül megváltoztatásra



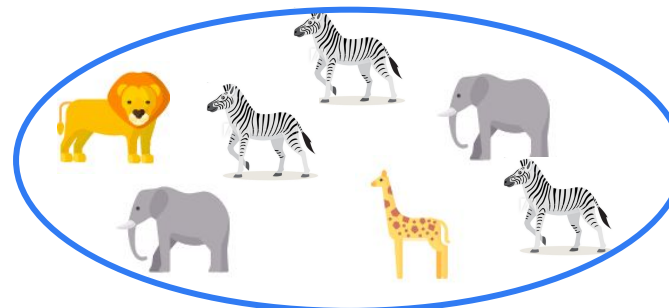
Szekvencia & Helyszín információk

Tipikus minták képsorozatokban

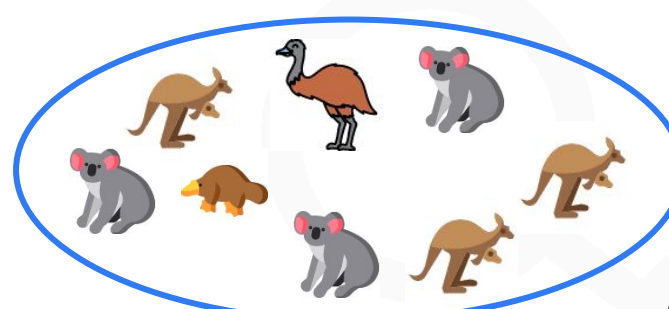
- Tipikusan egy állatfaj per sorozat
- A sorozatok végein nincs állat a képen
- A legmagabiztosabb predikció jól reprezentálja az egész sorozatot



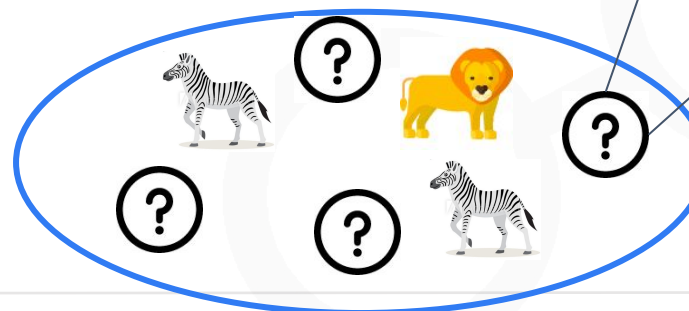
Train location 1



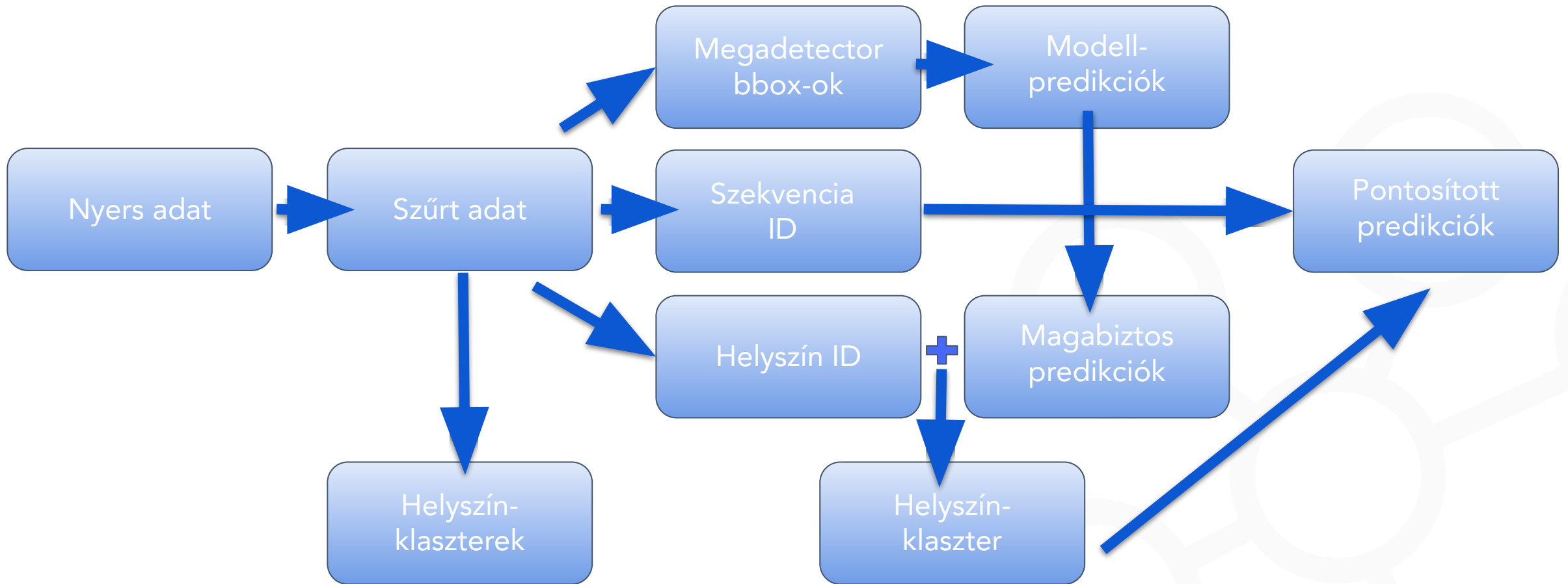
Train location 2



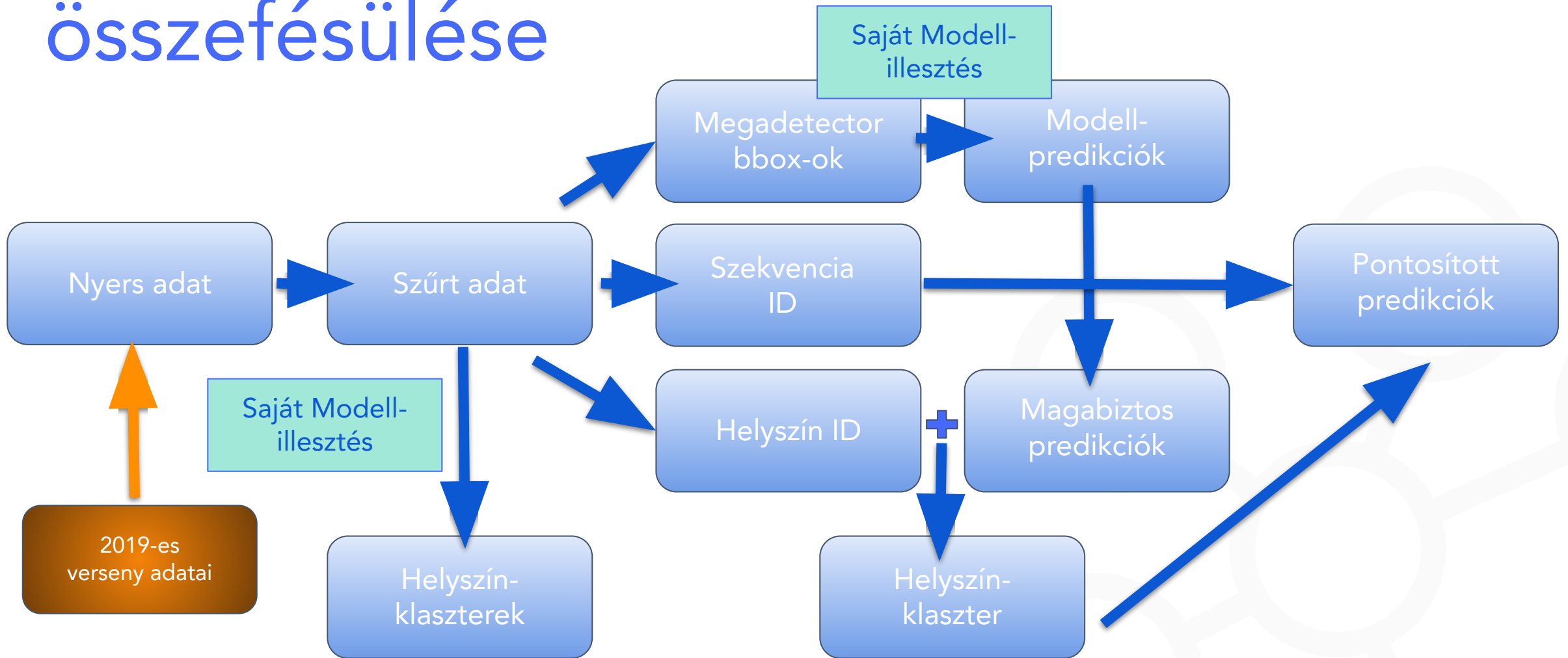
Test location 1



A különböző információforrások összefésülése



A különböző információforrások összefésülése

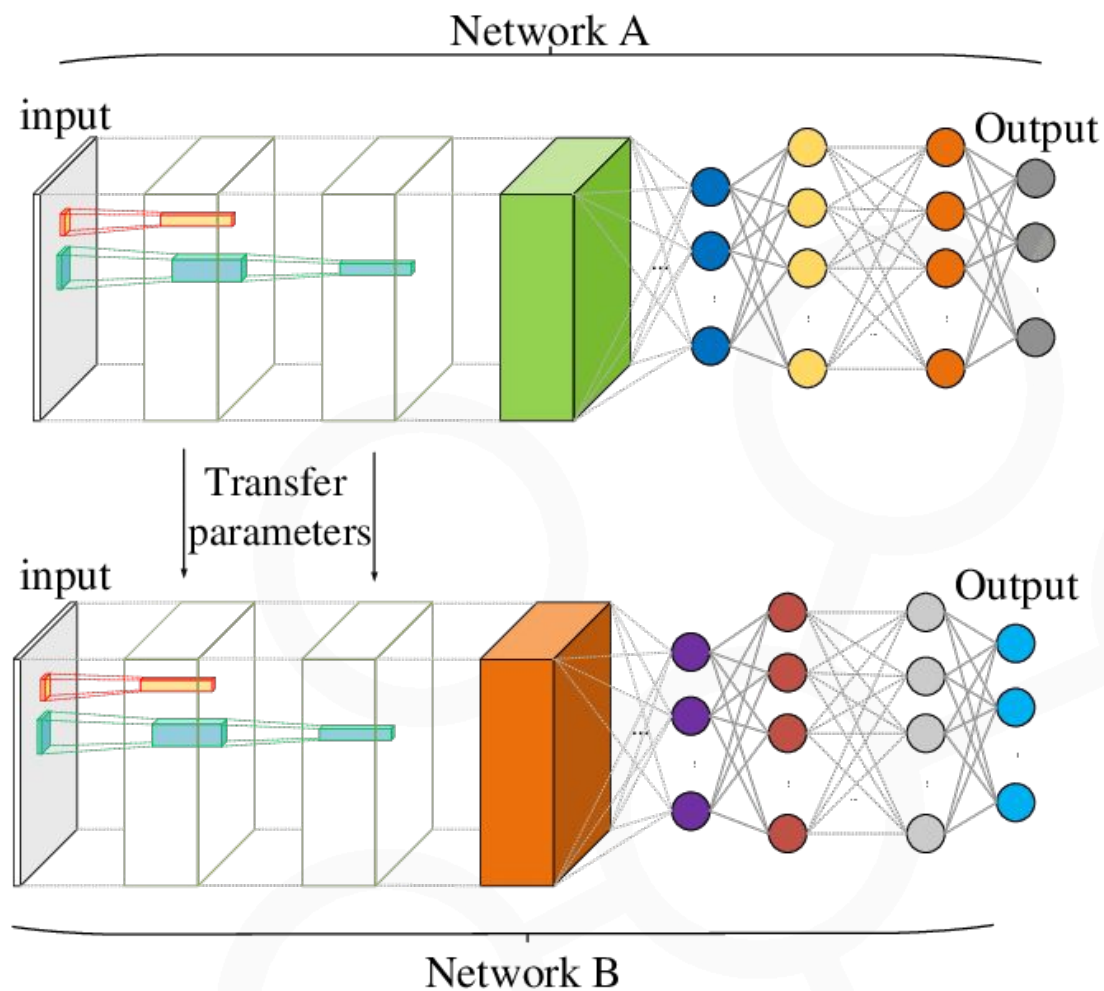


Transfer learning:

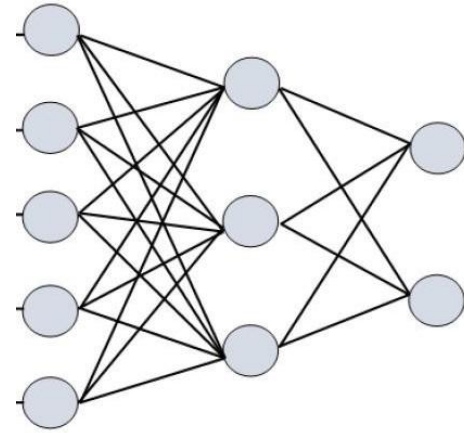
A transfer learning egy előre betanított modellt használ alapul, és sok esetben csak az utolsó réteget tanítják újra, adott feladatnak megfelelően.

Ennek nagy előnye:

- nem kell egyedi architektúrát fejleszteni
- nem kell annyi ideig trainelni



Architektúra keresés:



Minta méret: 8400 kép

	resnet18	resnet34	resnet50	resnet101	resnet152	vgg16	vgg19	densenet121	densenet169	densenet201	densenet161
batch size	64	64	64	64	32	64	64	64	64	32	32
epoch time	2:30	2:40	2:49	3:20	4:44	2:56	3:10	2:46	2:53	3:40	4:57
GPU memory	2100 MB	2500 MB	6600 MB	9400 MB	6700 MB	8700 MB	9600 MB	9300 MB	11050 MB	7500 MB	8800 MB
train loss	0.856132	0.829578	0.665367	0.627351	0.820779	0.887325	0.920419	0.70819	0.521828	0.561023	0.678316
valid loss	0.903686	0.866548	0.773739	0.736688	0.826549	0.851915	0.884898	0.776619	0.729133	0.696411	0.736785
error rate	0.233234	0.223145	0.199407	0.190504	0.223145	0.221958	0.219585	0.200593	0.179228	0.166172	0.188724
learning rate	0.00525	0.01318	0.00302	0.00251	0.00912	0.00209	0.00174	0.00145	0.00302	0.001445	0.005248

Ensemble:

Eltérő modellek (és architektúrák), amiknek mások az erősségeik és gyengeségeik (állatok, növényzet, megvilágítás, stb.)

Bayes-i paraméter kereséssel megkerestük a modellek súlyait, és ez alapján készült el a végső predikció.

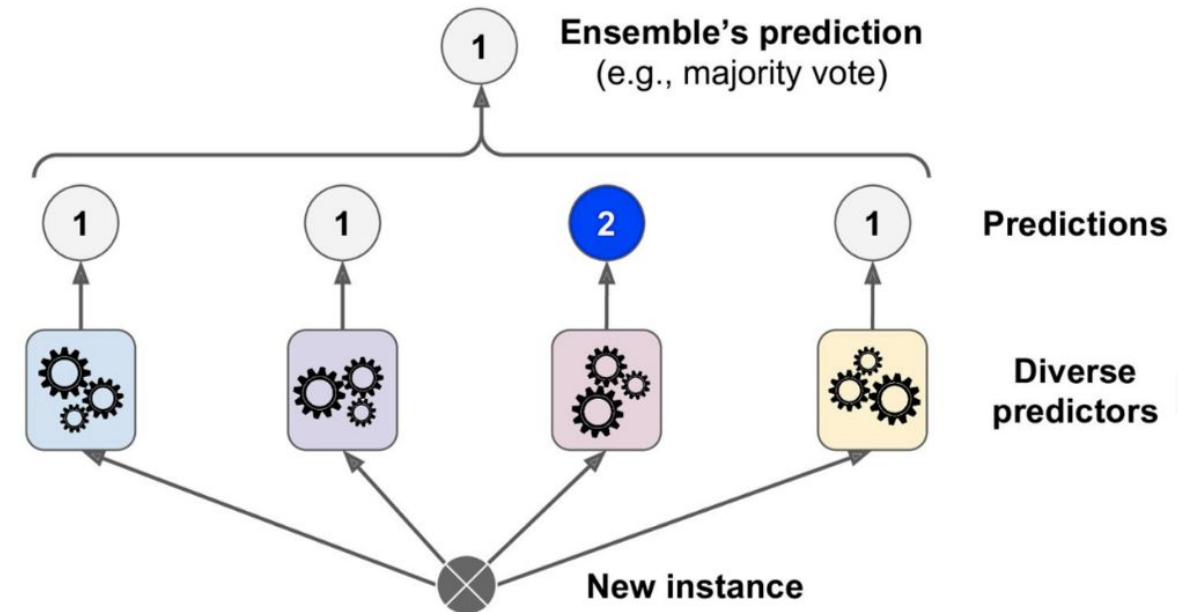


Figure 7-2. Hard voting classifier predictions

Más csapatok:



	1. helyezett	2. helyezett	3. helyezett	4. helyezett	Starschema
Publikus pontszám	0.953	0.93	0.845	0.86	0.771
Privát pontszám	0.906	0.902	0.848	0.822	0.778
Architektúra	Efficientnet-b6 SeResNeXt101	EfficientNetB4	EfficientNet, ResNet152, NTS with ResNet50, NTS with ResNext50	InceptionV3	Resnet34. Resnet50, vgg16, densenet121
Előző évi verseny képei		X			X
Dataset tisztítása		X	X		X
Data augmentation	X	X	X	X	X
Megadetector	X	X	X	X	X
Kivágott képen tanítás	X	X	X	X	X
Szekvencia (kép sorozat)	X	X	X	X	X
Lokáció klasszifikáció	X	X	X	X	X
Class imbalance	speciális mintavételezés		külön model	speciális mintavételezés	külön model
Ensemble	X	X	X		X

Konklúzió:

- Az adattisztítás (a képeknél is) fontos - érdemes kiszűrni a rossz minőségű képeket
- Új adat generálása: Object Detectionnel megkeresni az állatokat és utána klasszifikálni
- Egyszerű szabályok használata: szekvencia ID felhasználása
- Használjunk Ensemble learninget: A különböző modellek tranelése eltérő adatokon és kombináljuk ezeket a modelleket



STARSCHEMA

Köszönjük a figyelmet!

16 MPH



Képek előfeldolgozása

Hisztogram-kiegyenlítés

- A kép kontrasztjának megnövelése a színek komponensek eloszlásának uniformizálásával
- Adaptív változat: minden pixel a saját környezete alapján kerül megváltoztatásra

