

A GÉPI TANULÁSI MODELLEKKEL TÖRTÉNŐ ELŐREJELZÉSEK HATÉKONYSÁGÁT BEFOLYÁSOLÓ TÉNYEZŐK

DR. BOGNÁR LÁSZLÓ

DR. FAUSZT TIBOR

NAGY GÁBOR ZSOLT

LŐRINCZI KONRÁD



MOODLEMOOT
HU 2021

MINTA-MÁTRIX ÉS CÉL-VEKTOR

CORE INDIKÁTOROK

AZ INDIKÁTOROK SZÁMÍTÁSA A LOG-BEJEGYZÉSEKBŐL

MINTÁK – INDIKÁTOROK –CÉL

Indicators ?

No selection

Search

Analysis interval ?

Quiz cognitive X_1

Contexts ?

SCORM cognitive

SCORM social

Survey cognitive

Survey social

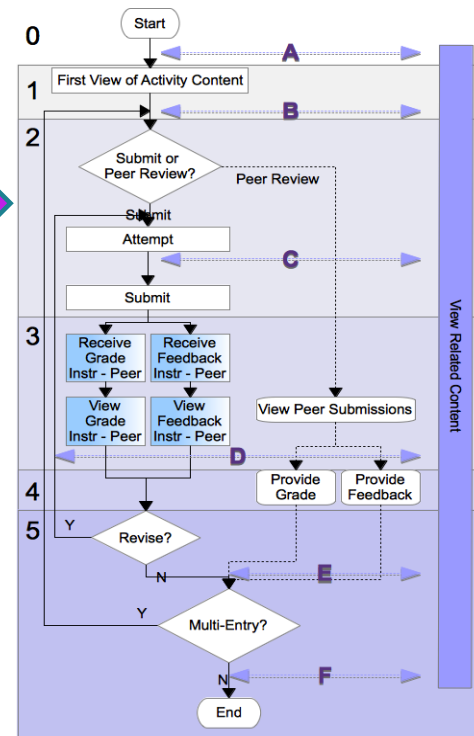
Predictions ?

processor

URL cognitive

URL social

Wiki cognitive



X
mátrix

Y
vektor

Minta	Hallgató	Indikátorok						Cél Y
		X_1	...	X_i	...	X_p		
1	1							0
2	...							1
...	...							...
Minta i	i							1
...	...							...
Minta n	n							1

IDŐOSZTÁSOS MÓDSZER A MOODLE-BAN

Minta	Hallgató	Indikátorok												Idő- osztás	Cél Y	
		Időosztás indikátorok					Core indikátorok									
		s ₁	...	s _k	...	s _c	x ₁	x ₁ Átlag	...	x _i	x _i Átlag	...	x _p			x _p Átlag
					X											
1	1	1		0		0									1	0
...	...	...		...		...										...
...	n	1		0		0										1
...	1	0		0		0									...	0
...	...	...		...		...										...
...	n	0		0		0										1
(k-1)n+1	1	0		1		0									k	0
...				1												...
...				1												...
Minta i	→	0		1		0										1
...	1	0		0		0									...	0
...	...	...		...		...										...
...	n	0		0		0										1
...	1	0		0		1									c	0
...	...	...		...		...										...
m = cn	n	0		0		1										1

X mátrix

Y vektor

A MATEMATIKAI MODELL

$$Q = (Y(\text{tényleges}) - Y(\text{a modell által előre jelzett}))^2 = \text{Minimum}$$

Indikátorok

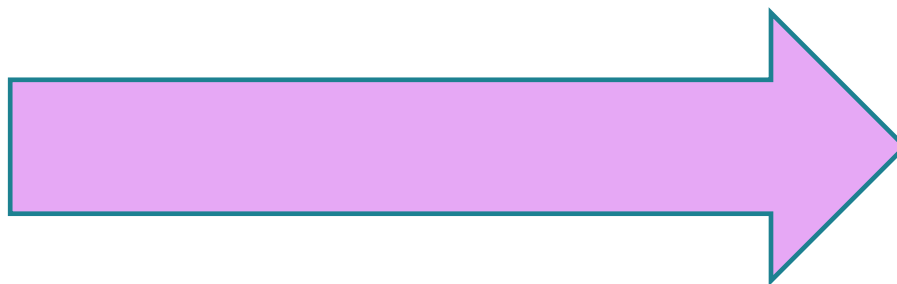
X_1

X_2

...

X_p

$$Y = f(X_1, \dots, X_j, \dots, X_p) + \varepsilon$$



Két opció a Moodle-ban:

- Logisztikus regresszió
- Neurális háló

Tanulási cél

$Y = 1$ ha a hallgató sikertelen

$Y = 0$ ha a hallgató sikeres

A MODELL JÓSÁGÁNAK MÉRŐSZÁMAI

TÉVESZTÉSI MÁTRIX

ELŐRE JELZETT CÉL

0

1

TÉNYLEGES CÉL
0
1

0	TN True Negative	FP False Positive
1	FN False Negative	TP True Positive

Accuracy = $(TP + TN)/n$; **Error** = $1 - \text{Accuracy}$

Recall = $TP/(TP + FN)$; a ténylegesen sikertelen hallgatók közül azoknak az aránya, akiket a modell is sikertelennek jósolt

Fallout = $FP/(FP + TN)$; a kurzust ténylegesen sikeresen teljesítő hallgatók közül azoknak az aránya, akiket a modell félrediagnosticsztizált, azaz sikertelennek jósolt

Precision = $TP/(TP + FP)$; azok közül a hallgatók közül, akiket a modell sikertelennek jósolt, azoknak az aránya, akik ténylegesen sikertelenek lettek

F1 Score [0-1]

nMCC [0-1]

A MODELL JÓSÁGÁT LEGINKÁBB BEFOLYÁSOLÓ TÉNYEZŐK I.

- 1. Hány és milyen indikátort használunk a modellben? Milyen Moodle-erőforrásokat használunk? Hogyan számítjuk az indikátorok értékét?
- 2. Nappali vagy levelező a kurzus? Hány félévi tesztből szerzik a hallgatók a vizsgajegyüket?
- 3. Mennyire nehéz a kurzus? Mennyi a minimálisan elérendő pontszám?
- 4. Hány hallgató van a kurzusban? Hány időosztást használunk? Mennyi a sorok és az oszlopok aránya az X mátrixban?
- 5. Hogyan tanítjuk a modellünket, milyen időintervallumokat használunk a betanításhoz és az előrejelzéshez?

A MODELL JÓSÁGÁT LEGINKÁBB BEFOLYÁSOLÓ TÉNYEZŐK II.

- 6. Milyen a Moodle-on belül és kívül zajló tanulási tevékenységek aránya?
- 7. Milyen indikátorokat használunk? A hallgatókra vagy az ő tevékenységeikre vonatkozókat? Mennyire szorosan kapcsolódnak az indikátorok értékei a sikerességhez?
- 8. Milyen matematikai módszerrel tanítjuk a modellt (logisztikus regresszió, neurális háló stb.)?
- 9. Mi a kurzus tartalma (pl. Matematika kontra esszéírás)?
- 10. Kurzus szintű vagy portál szintű modellt építünk?

TANULÁSI FORRÁSOK

- Lecture Notes
- Exercise Books
- Lecture Videos
- Learning the Concepts Quizzes
- Chapter Review Quizzes
- Minitab Videos



1. HÁNY ÉS MILYEN INDIKÁTORT HASZNÁLUNK A MODELLBEN?
MILYEN MOODLE-ERŐFORRÁSOKAT HASZNÁLUNK? HOGYAN
SZÁMÍTJUK AZ INDIKÁTOROK ÉRTÉKÉT?



AZ ELEMZÉSEKBEN HASZNÁLT CORE INDIKÁTOROK

- File Cognitive (1 indikátor a teljes kurzusra)
- Book Cognitive (1 indikátor a teljes kurzusra)
- Page Cognitive (1 indikátor a teljes kurzusra)
- Quiz Cognitive (1 indikátor a teljes kurzusra)

Az összes indikátor darabszáma = 4

AZ ELEMZÉSEKBEN HASZNÁLT SAJÁT INDIKÁTOROK

- Lecture Notes; megtekintés száma (7 indikátor; fejezetenként 1 db)
- Exercise Books; megtekintés száma (7 indikátor; fejezetenként 1 db)
- Lecture Videos; megtekintés száma (7 indikátor; fejezetenként 1 db)
- Minitab videos; megtekintés száma (8 indikátor; Minitab fejezetenként 1 db)
- Quiz Attempts; darabszám (7 vagy 14 indikátor; 1 vagy 2 db fejezetenként)
- Quiz Max Grades; pont (7 vagy 14 indikátor; 1 vagy 2 db fejezetenként)

Az összes indikátor darabszáma = 43 vagy 57

MODELLEK MOODLE CORE INDIKÁTOROKKAL

Nappali kurzus
Sikertelen = 48%

Modell 0 File Cognitive 1 indikátor	Modell 1 Page Cognitive 1 indikátor	Modell 2 Book Cognitive 1 indikátor	
Modell 4 File Cognitive + Page Cognitive + Book Cognitive 3 indikátor			Modell 3 Quiz Cognitive 1 indikátor
Modell 5 File Cognitive + Quiz Cognitive + Page Cognitive + Book Cognitive 4 indikátor			

MODELLEK SAJÁT INDIKÁTOROKKAL

Modell 6 Lecture Notes 7 indikátor	Modell 7 Exercise Books 7 indikátor	Modell 8 Lecture videos 7 indikátor	Modell 9 Minitab videos 8 indikátor		
Modell 12 Lecture Notes + Exercise Books 14 indikátor		Modell 13 Lecture videos + Minitab videos 15 indikátor		Modell 10 Quiz Attempts 14 indikátor	Modell 11 Quiz Max Grades 14 indikátor
Modell 14 Lecture Notes + Exercise Books + Lecture videos + Minitab videos 29 indikátor				Modell 15 Quiz Attempts + Quiz Max Grades 28 indikátor	
Modell 16 Lecture Notes + Exercise Books + Lecture videos + Minitab videos + Quiz Attempts + Quiz Max Grades 57 indikátor					

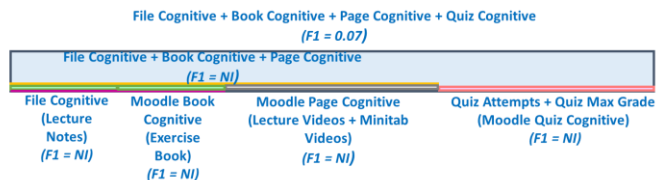
A MODELLEK JÓSÁGÁNAK ÖSSZEHASONLÍTÁSA

4 Moodle Core indikátor

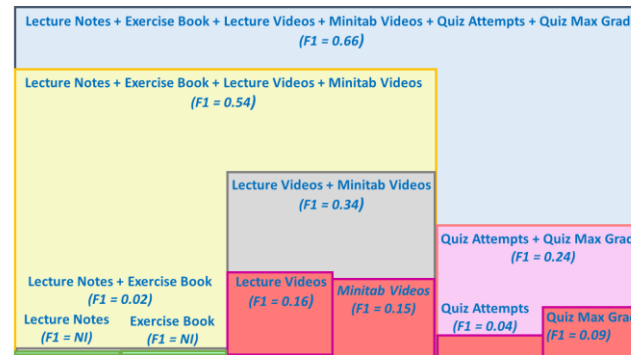
43 saját indikátor
(14 Quiz indikátor)

57 saját indikátor
(28 Quiz indikátor)

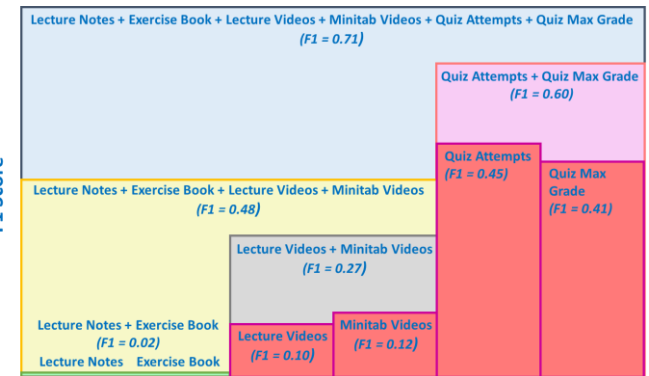
F1 Score



F1 score



F1 Score

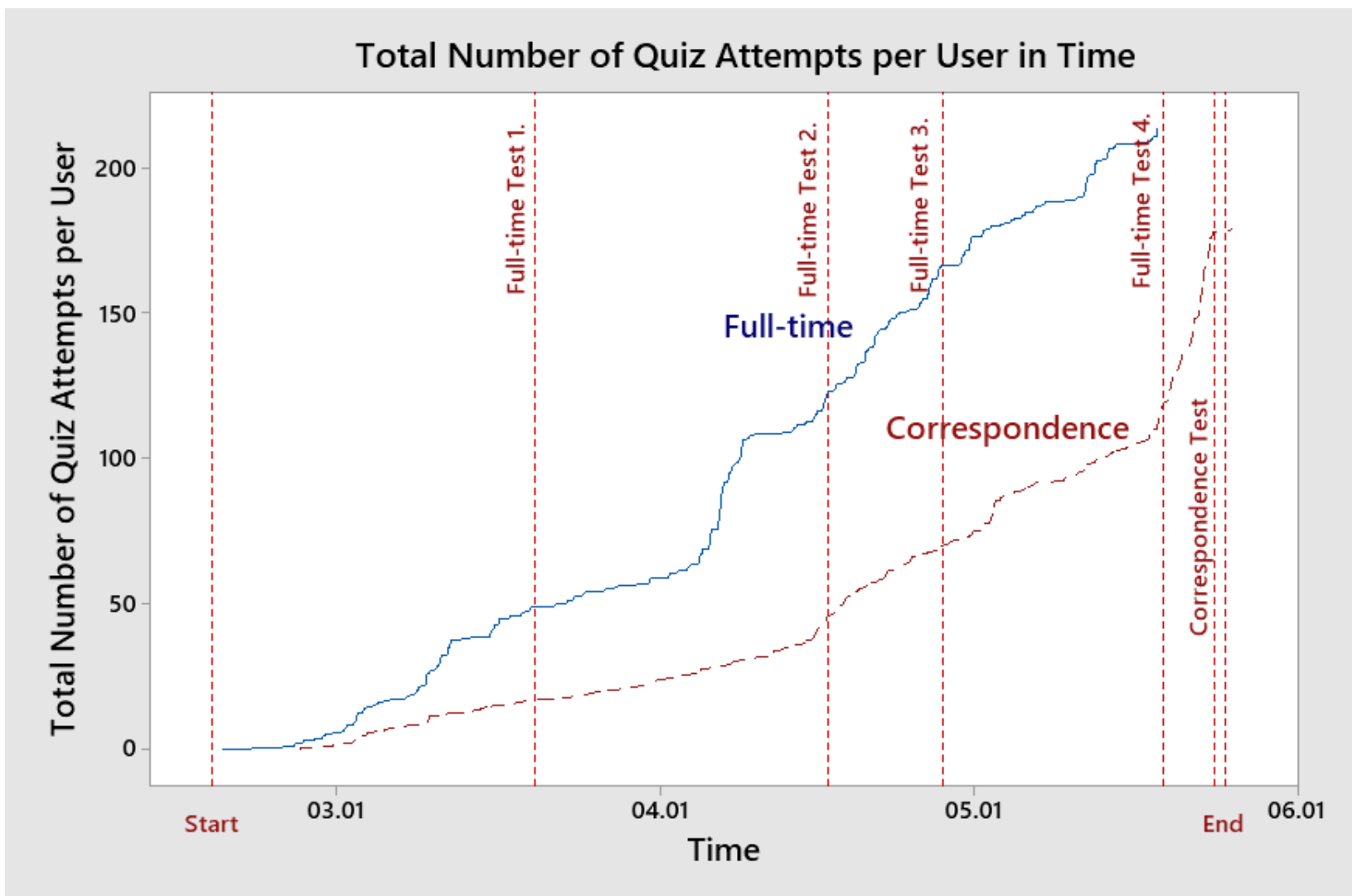




2. NAPPALI VAGY LEVELEZŐ A KURZUS? HÁNY FÉLÉVI
TESZTBŐL SZERZIK A HALLGATÓK A VIZSGAJEGYÜKET?



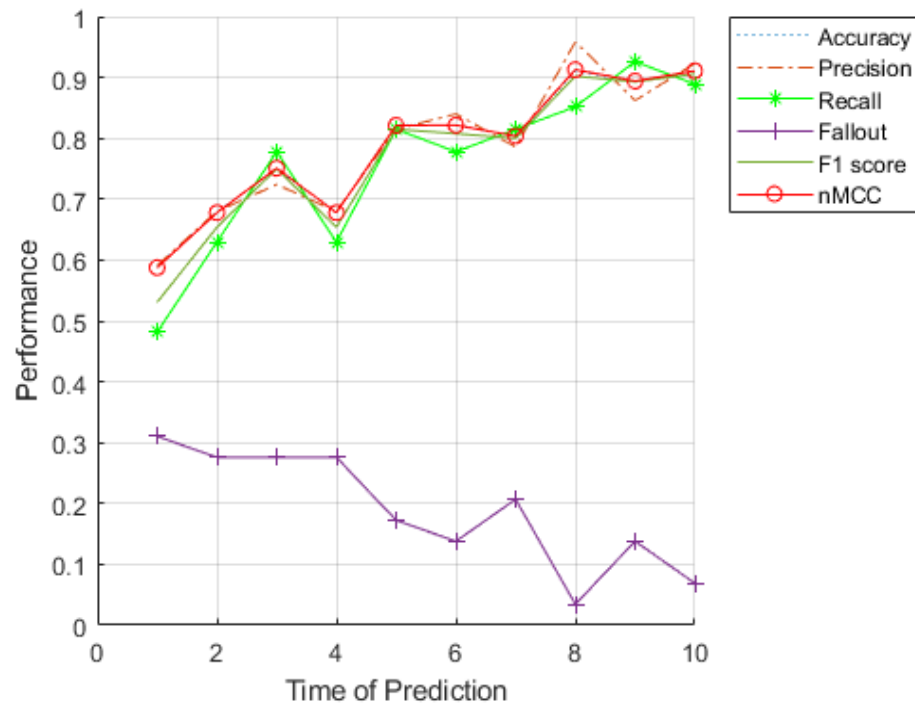
TANULÁSI AKTIVITÁS AZ IDŐBEN



A MODELEK JÓSÁGÁNAK ÖSSZEHAISONLÍTÁSA

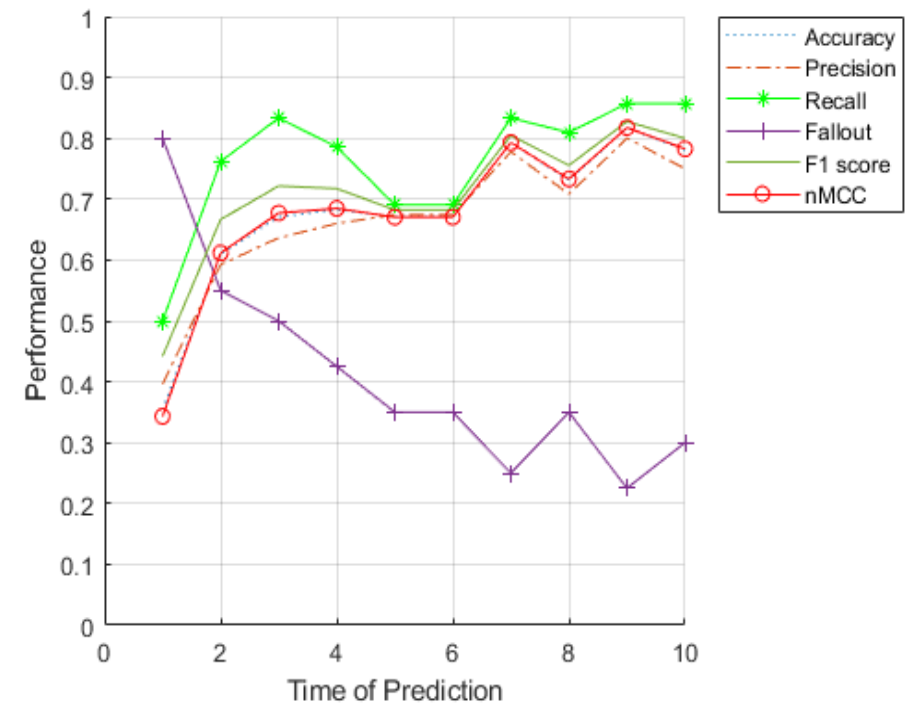
Nappali kurzus

Model Performance Curves in Time



Levelező kurzus

Model Performance Curves in Time





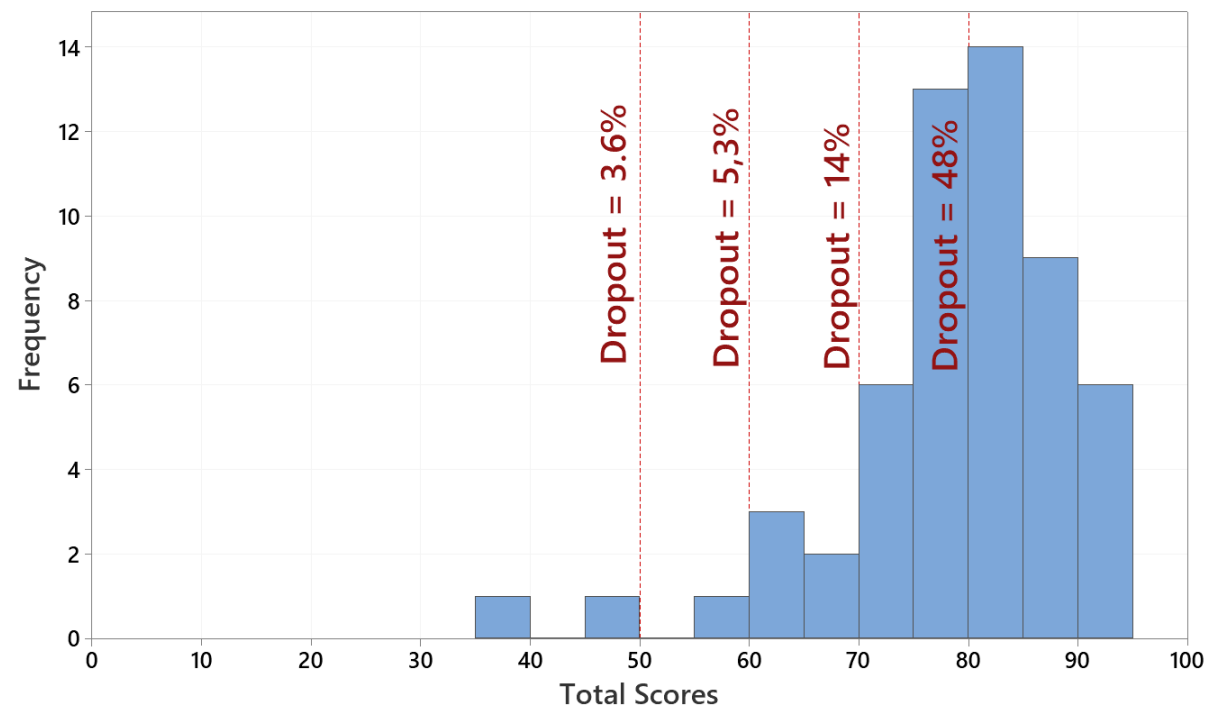
3. MENNYIRE NEHÉZ A KURZUS? MENNYI A MINIMÁLISAN
ELÉRENDŐ PONTSZÁM?



A SIKERTELEN HALLGATÓK ARÁNYA

Nappali kurzus

Histogram of Total Scores
Full-time Course



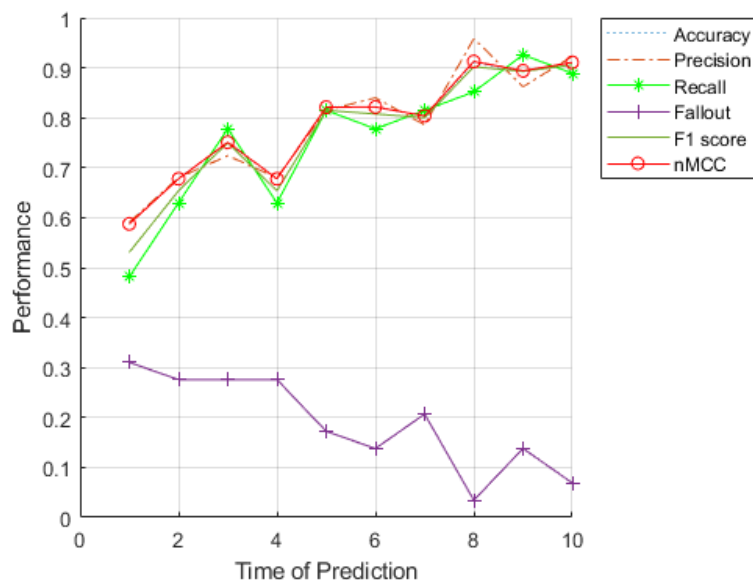
A MODELLEK JÓSÁGÁNAK ÖSSZEHAISONLÍTÁSA

Minimum pontszám = 80
Sikertelen = 48%

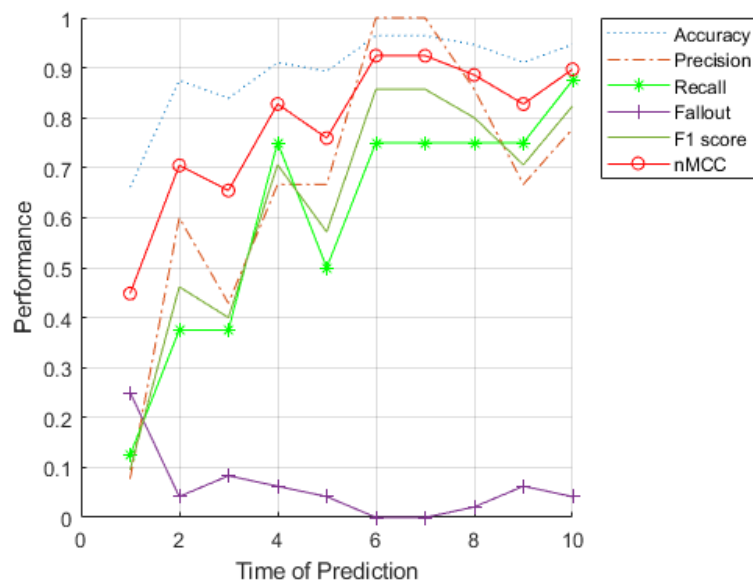
Minimum pontszám = 70
Sikertelen = 14%

Minimum pontszám = 60
Sikertelen = 5.3%

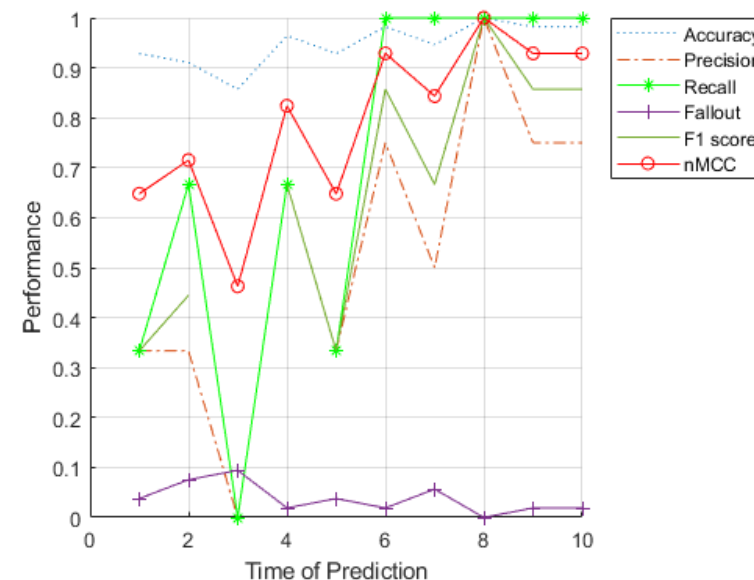
Model Performance Curves in Time



Model Performance Curves in Time



Model Performance Curves in Time



Nappali kurzus; 57 saját indikátor

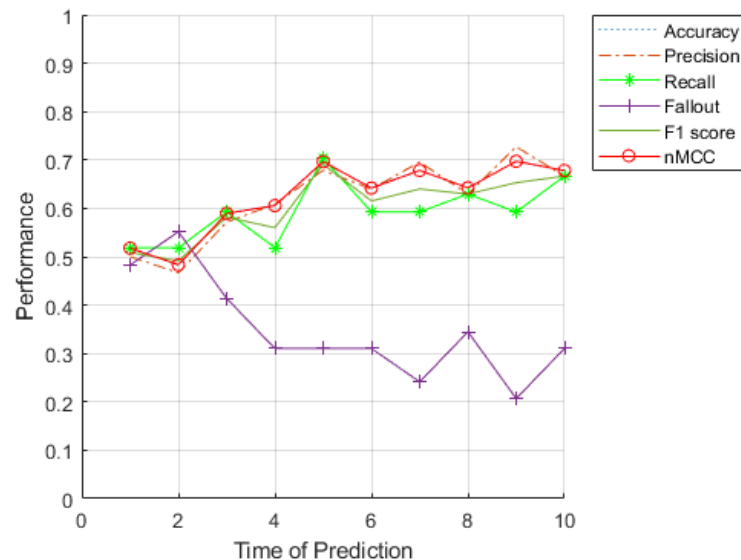
A MODELLEK JÓSÁGÁNAK ÖSSZEHAISONLÍTÁSA

Minimum pontszám = 80
Sikertelen = 48%

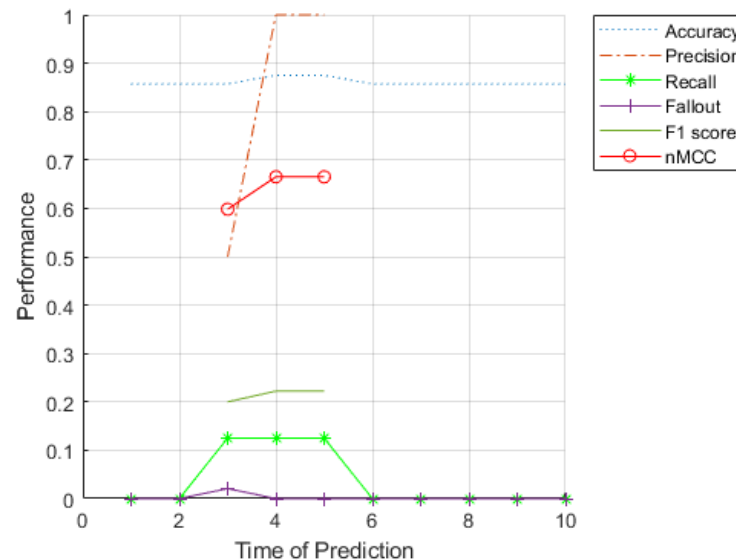
Minimum pontszám = 70
Sikertelen = 14%

Minimum pontszám = 60
Sikertelen = 5.3%

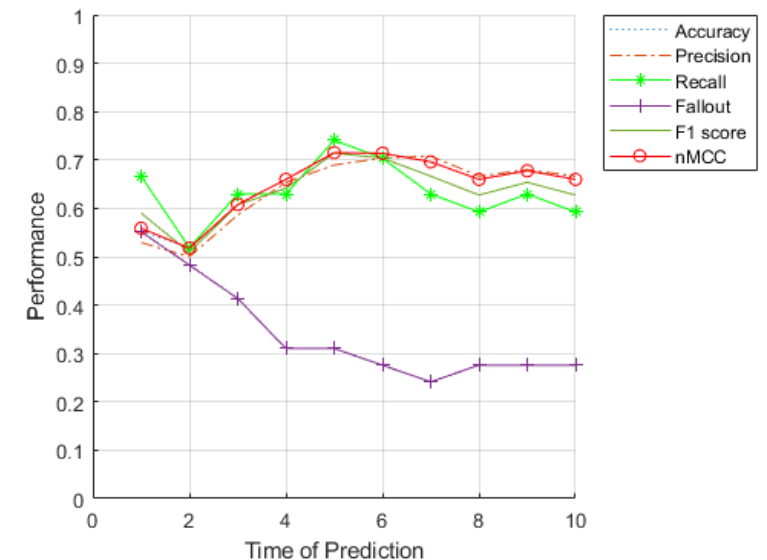
Model Performance Curves in Time



Model Performance Curves in Time



Model Performance Curves in Time



Nappali kurzus; 4 Moodle Core indikátor



8. Milyen matematikai módszerrel tanítjuk a modellt (logisztikus regresszió, neurális háló stb.)?

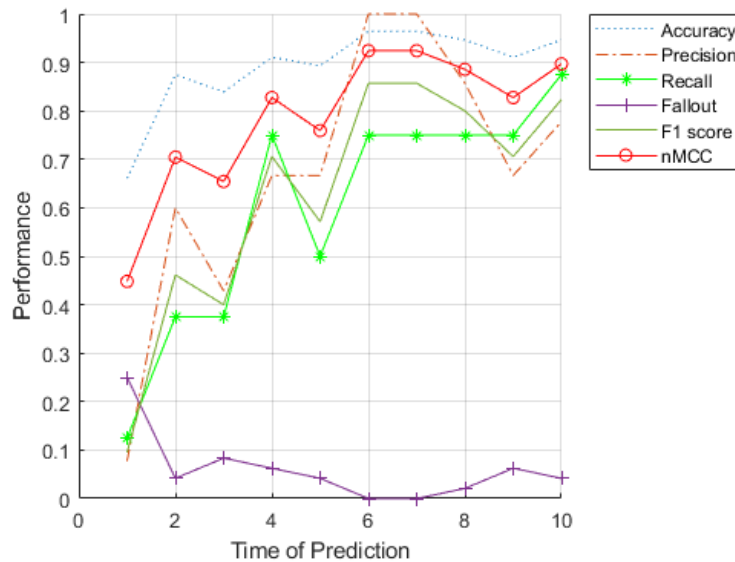


A MODELEK JÓSÁGÁNAK ÖSSZEHAISONLÍTÁSA

Logisztikus regresszió

	0	1
0	445	25
1	28	52

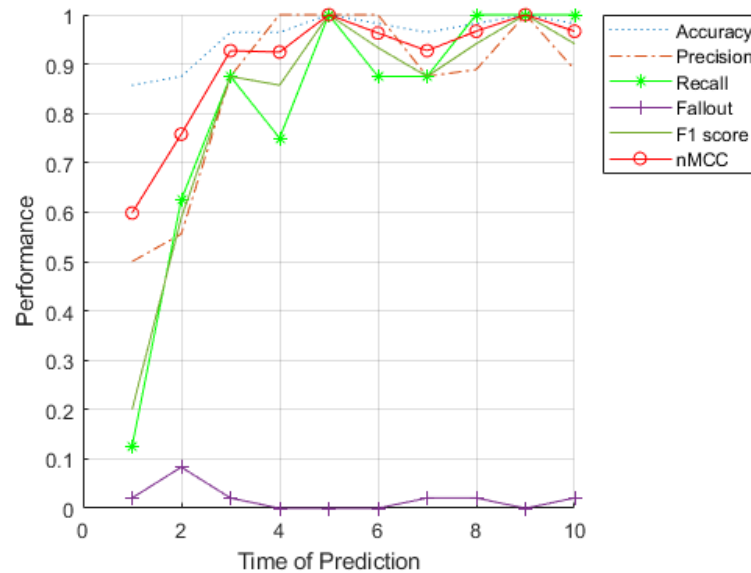
Model Performance Curves in Time



Klasszifikációs fa

	0	1
0	466	14
1	14	66

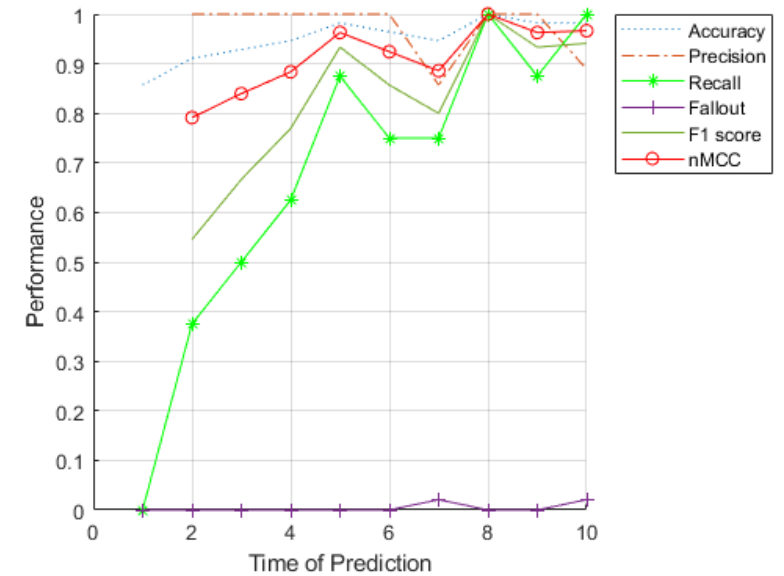
Model Performance Curves in Time



Support Vector Machine

	0	1
0	478	2
1	30	50

Model Performance Curves in Time



Nappali kurzus; Sikertelen = 14%; 43 saját indikátor

A MODELEK JÓSÁGÁNAK ÖSSZEHAISONLÍTÁSA

Training Confusion Matrix

Output Class	1	2	
	330 84.2%	24 6.1%	93.2% 6.8%
2	4 1.0%	34 8.7%	89.5% 10.5%
	98.8% 1.2%	58.6% 41.4%	92.9% 7.1%
		Target Class	

Validation Confusion Matrix

Output Class	1	2	
	73 86.9%	5 6.0%	93.6% 6.4%
2	1 1.2%	5 6.0%	83.3% 16.7%
	98.6% 1.4%	50.0% 50.0%	92.9% 7.1%
		Target Class	

Test Confusion Matrix

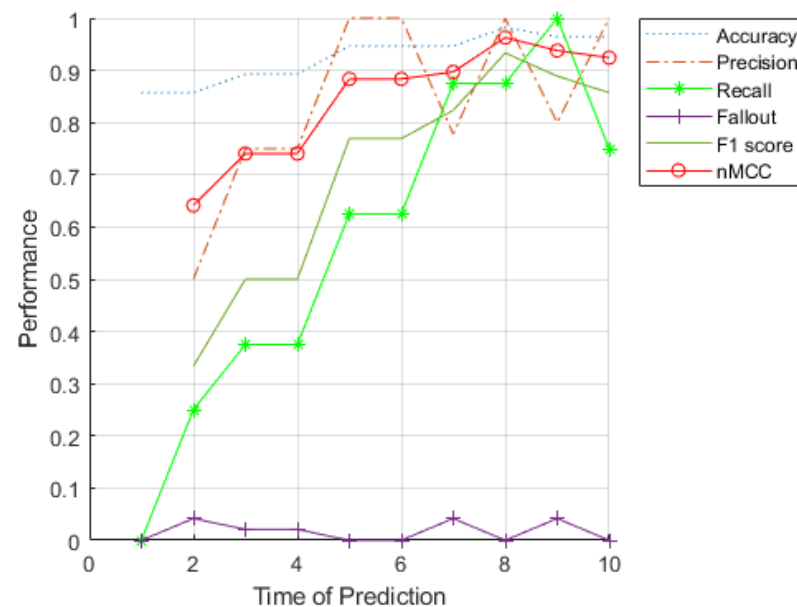
Output Class	1	2	
	69 82.1%	5 6.0%	93.2% 6.8%
2	3 3.6%	7 8.3%	70.0% 30.0%
	95.8% 4.2%	58.3% 41.7%	90.5% 9.5%
		Target Class	

All Confusion Matrix

Output Class	1	2	
	472 84.3%	34 6.1%	93.3% 6.7%
2	8 1.4%	46 8.2%	85.2% 14.8%
	98.3% 1.7%	57.5% 42.5%	92.5% 7.5%
		Target Class	

Neurális háló

Model Performance Curves in Time



Nappali kurzus;
Sikertelen = 14%;
43 saját indikátor



ÖSSZEGZÉS JAVASLATOK



- Az előrejelzésre használt kurzus jellemzőinek szigorúan meg kell egyezni a modell betanításához használt kurzus jellemzőivel.
- Sokkal több matematikai részletet kellene közölni a professzionális Moodle-felhasználók, matematikusok, statisztikusok számára.
- Sokkal több felhasználóbarát eszközre lenne szükség az „átlagos” Moodle felhasználók számára a modellek felépítéséhez és értékeléséhez, valamint az előrejelzések megbízhatóságának elemzéséhez.
- Szekció- és erőforrásszintű indikátorokat kellene beépíteni a Moodle core-ba, hogy ezek kurzuson belüli számát növelni lehessen.
- Az előrejelzések időbeli részletei fontosak lennének.
- Nemcsak a portál adminisztrátoroknak, hanem a tanári jogosultsággal rendelkezőknek is meg kellene kapniuk a jogot ezen eszközök használatára.

Köszönjük a figyelmet!

Dr. **Bognár** László

bognar.laszlo@uniduna.hu

Dr. **Fauszt** Tibor

fauszt.tibor@uni-bge.hu

Nagy Gábor Zsolt

gnagy@eduroll.eu

Lőrinczi Konrád

klor@uniduna.hu



MOODLEMOOT
HU 2021