



Tantárgy kód

**BMETE95AM36**

**Tantárgy azonosító adatok**

1.	A tárgy címe	<b>Bevezetés az adattudományba 1</b>							
2.	A tárgy angol címe	<b>Introduction to data science 1</b>							
3.	Heti óraszámok (ea + gy + lab) és a félévvégi követelmény típusa	<b>3</b>	+	<b>0</b>	+	<b>1</b>	<b>v</b>	Kredit	<b>4</b>
4.	Ajánlott/kötelező előtanulmányi rend								
	vagy	Tantárgy kód 1	Rövid cím 1	Tantárgy kód 2	Rövid cím 2	Tantárgy kód 3	Rövid cím 3		
	4.1	<b>BMETE91AM43</b>	<b>Informatika2</b>	<b>BMETE95AM31</b>	<b>Statisztika1</b>				
	4.2								
	4.3								
5.	Kizáró tantárgyak								
6.	A tantárgy felelős tanszéke	<b>Sztochasztika Tanszék</b>							
7.	A tantárgy felelős oktatója	<b>Dr. Simon Károly</b>	beosztása	<b>egyetemi tanár</b>					

**Akkreditációs adatok**

8.	Akkreditációra benyújtás időpontja	<b>2015.02.16.</b>	Akkreditációs bizottsági döntés időpontja	<b>2016.04.18.</b>
----	------------------------------------	--------------------	---	--------------------

**Megjegyzések**

Csak az űrlap fehéren hagyott mezőibe írjunk és a mezők között a **tabulátor** billentyűvel haladjunk! Ha egy kitöltött mezőből tabulátor billentyűvel lépünk ki, több más mező értéke automatikusan megváltozhat. Egy adott mezőre lépve, az állapotsorban megjelenő rövid, ill. az F1 gomb megnyomásakor kapható hosszabb leírás ad segítséget a kitöltéshez. A *tantárgy kódot* és a *tárgy rövid címét* a dékáni hivatal adja.

**1-2. sorok:** A *tárgy címének* (max. 85 karakter) célszerű legalább egy karakterben különböznie minden más, Neptunban regisztrált tárgy címétől.

**3. sor:** A *követelmény* előadás+gyakorlat+labor formátumú, az *utolsó mező* a félév végi számonkérés típusa (v,f,a vagy s, részletes információ F1). A *kredit* megadásánál ügyelni kell arra, hogy az alább részletezett, a *tantárgy elvégzéséhez átlagosan szükséges tanulmányi munka* mennyiségével összhangban legyen (összes óraszám = kredit\*30 óra).

**4. sor:** Legfeljebb 3, már korábban hallgatott tárgy adható meg a 4.1 sorban. A 4.2 és 4.3 sorok *vagylagos* lehetőségek megadására szolgálnak, például abban az esetben, ha az egyik tárgynak korábban oktatott változatai is megfelelőek. **5. sor:** A *kizáró tantárgyaknál* azokat a tárgyakat kell felsorolni, amelyek tematikái a most akkreditálandó tárggyal 75% vagy annál nagyobb átfedést mutatnak.

**6-7. sorok:** A felelős tanszék és oktató hatáskörét, ill. kijelölésének feltételeit a *Képzési Kódex 2010* c. dokumentum 4.§-a tartalmazza.

<b>Tematika</b>			
9.	A tantárgy az alábbi témakörök ismeretére épít <b>Statisztika, inearis algebra, Python programozás</b>		
10.	A tantárgy szerepe a képzés céljának megvalósításában (szak, kötelező, kötelezően választható, szabadon választható) <b>TTK Matematika (BSc) képzés Adattudományi sávjának kötelező tárgya.</b>		
11.	<p>A tantárgy részletes tematikája</p> <p><b>A tárgy célja az adattudomány alapfogalmainak a korábban megszerzett matematikai ismeretekre épülő, gyakorlati megközelítésű megismertetése. A hallgatók a kezdetektől teljes, a gyakorlati életből vett valós alkalmazási példákon keresztül az ismereteket megtapasztalva, egyfajta spirál mentén egyre mélyebbre haladva precíz elméleti és egyúttal praktikus gyakorlati ismeretekhez jutnak. Az elméleti ismeretek gerincét a gépi tanulás algoritmusai adják, a gyakorlati feladatok építenek a Python nyelv ismeretére.</b></p> <p><b>Előadás: Történet, példák, esettanulmányok, az adattudományba sorolható diszciplínák. Ellenőrzött tanulás - Lineáris Modellek + modell validálás. Legkisebb négyzetek módszere. Lineáris Regresszió. Gradiens módszer, maximum-likelihood becslés. Polinomiális regresszió, logisztikus regresszió, Perceptron, Newton-módszer, Naive-Bayes. Általánosított lineáris modellek (Exponenciális család), tanulási/validációs/tesztelési halmaz, cross-validáció, Bias-Variance tradeoff, regularizáció, Precision-Recall, F1-score, ROC görbe. SVM, lineáris SVM, kernel trükk. Neurális hálók. Döntési fák. Véletlen erdők. Boosting. Nem Ellenőrzött tanulás. Klaszterezés. K-means klaszterezés. EM algoritmus. PCA, ICA. Nagyobb esettanulmányok, kitekintés.</b></p> <p><b>Gyakorlat: Az adatmanipulálás, prediktív analízis, megjelenítés lépései valódi adatokkal (pl. kaggle) elsősorban Python-csomagok (pandas, scikit-learn, matplotlib, ggplot) és R használatával.</b></p>		
12.	Követelmények, az osztályzat (aláírás) kialakításának módja		
	szorgalmi időszakban	<b>házi feladatok, részvétel a tanórákon</b>	vizsga-időszakban <b>vizsga</b>
13.	Pótlási lehetőségek <b>TVSZ szerint</b>		
14.	Konzultációs lehetőségek <b>TVSZ szerint</b>		
15.	Jegyzet, tankönyv, felhasználható irodalom <b>legfrissebb online jegyzetek</b>		

16.	A tantárgy elvégzéséhez átlagosan szükséges tanulmányi munka mennyisége órákban (a teljes szemeszterre számítva)		
	16.1	Kontakt óra	56
	16.2	Félévközi felkészülés órákra	14
	16.3	Felkészülés zárthelyire	0
	16.4	Zárthelyik megírása	0
	16.5	Házi feladat elkészítése	28
	16.6	Kijelölt írásos tananyag elsajátítása (beszámoló)	0
	16.7	Egyéb elfoglaltság	0
	16.8	Vizsgafelkészülés	22
	16.9	<b>Összesen</b>	<b>106</b>
17.	Ellenőrző adat		<b>Kredit * 30</b> <b>120</b>

A tantárgy tematikáját kidolgozta			
18.	Név	beosztás	Munkahely (tanszék, kutatóintézet stb.)
	<b>Dr. Simon Károly</b>	<b>egyetemi tanár</b>	<b>Sztochasztika Tanszék</b>

A tanszékvezető		
19.	Neve	aláírása
	<b>Dr. Simon Károly</b>	

**Megjegyzések**

**16.1 sor:** Értéke automatikusan kitöltődik az űrlap elektronikus változatában, a „Követelmény” címszónál megadott óraszám értékek alapján, az (előadás+gyakorlat+labor) \* (14 oktatási hét) formula szerint. **16.4 sor:** Értéke 0, ha a zárthelyik íratása kontakt órákon történik, egyébként pedig a minimálisan szükséges számú zárthelyi megírásához felhasználandó idő (a pót zárthelyik nélkül). **16.7 sor:** Az „Egyéb elfoglaltság” szöveg helyére a tevékenység konkrét megnevezését kell írni.

**17. sor:** Az itt szereplő értéknek és a **16.9 sorban** automatikusan megjelenő tanulmányi óraszám összegnek hozzávetőlegesen meg kell egyeznie! Tájékoztatásul azt vegyük figyelembe, hogy a hallgatók által egy szemeszterben átlagosan 30 kreditnyi munkamennyiséget kell teljesíteni, azaz a szorgalmi és vizsgaidőszak során elvárt terhelés összesen kb. 900 munkaóra.