



Tantárgy kód **BMETE91AM51**

| Tantárgy azonosító adatok |  |                                    |                    |                       |             |                |                          |
|---------------------------|--|------------------------------------|--------------------|-----------------------|-------------|----------------|--------------------------|
| 1.                        | A tárgy címe   | <b>Nyelvi adatok feldolgozása</b>  |                    |                       |             |                |                          |
| 2.                        | A tárgy angol címe   | <b>Natural Language Processing</b> |                    |                       |             |                |                          |
| 3.                        | Heti óraszámok (ea + gy + lab) és a félévvégi követelmény típusa | <b>2</b>                           | <b>+</b>           | <b>0</b>              | <b>+</b>    | <b>0</b>       | <b>v</b> Kredit <b>2</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>BMETE95AM36</b>                 | <b>BevAdattud1</b> |                       |             |                |                          |
|                           | 4.2  |                                    |                    |                       |             |                |                          |
|                           | 4.3  |                                    |                    |                       |             |                |                          |
| 5.                        | Kizáró tantárgyak  |                                    |                    |                       |             |                |                          |
| 6.                        | A tantárgy felelős tanszéke                                      | <b>Algebra Tanszék</b>             |                    |                       |             |                |                          |
| 7.                        | A tantárgy felelős oktatója                                      | <b>Dr. Kornai András</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<br><b>Python programnyelv, statisztika, lineáris algebra</b>  |                                       |                        |
| 10.             | A tantárgy szerepe a képzés céljának megvalósításában (szak, kötelező, kötelezően választható, szabadon választható)<br><b>TTK Matematika (BSc) képzés Adattudományi sávjának kötelezően választható tárgya.</b>   |                                       |                        |
| 11.             | A tantárgy részletes tematikája<br><b>A tárgy célja, hogy a nem numerikus, szöveges adatok (web, digitális könyvtárak, blogok) vagy a strukturálatlan (szabad szöveges) mezők nélkül nem értelmezhető strukturált, numerikus, képi adatok feldolgozásához szükséges nyelvi elemzési módszereket megismertesse a hallgatókkal.</b><br><br><b>Tematika: Adatgyűjtés. Végesállapotú technológiák. Környezetfüggetlen nyelvtanok. A szavak megszámlálása. Zipf törvényei, hatványtörvények. Indexépítés. A keresőmotorok alapjai. Amit a nyelvészetből tudni kell. A szavak osztályozása. Szótárépítés. Kollokációk, idiómák, többértelműség. Nyelvmodellezés. Súlyozott automaták, Markov modellek, rejtett Markov, n-gram. Helyesírás-ellenőrzés, nyelvtan-ellenőrzés. Beszédfelismerés, írásfelismerés, beszédkeltés. Névelemfelismerés. Funkcionális mondatelemzés. Mondat feletti egységek. Érzület-elemzés. Jelentésreprezentáció. Szójelentés, mondatjelentés, diskurzus-jelentés. Logikai modellek, vektormodellek. Gépi fordítás.</b> |                                       |                        |
| 12.             | Követelmények, az osztályzat (aláírás) kialakításának módja  |                                       |                        |
|                 | szorgalmi időszakban   | <b>programozási feladatok beadása</b> | vizsga-időszakban      |
|                 |  |                                       | <b>írásbeli vizsga</b> |
| 13.             | Pótlási lehetőségek<br><b>TVSZ szerint</b>   |                                       |                        |
| 14.             | Konzultációs lehetőségek<br><b>TVSZ szerint</b>  |                                       |                        |
| 15.             | Jegyzet, tankönyv, felhasználható irodalom<br><b>D Jurafsky, JH Martin: An Introduction to Natural Language Processing, Prentice Hall, 2009</b>  |                                       |                        |
|                 |  |                                       |                        |
|                 |  |                                       |                        |

|     |  |   |                                 |
|-----|--|---|---------------------------------|
| 16. | A tantárgy elvégzéséhez átlagosan szükséges tanulmányi munka mennyisége órákban (a teljes szemeszterre számítva) |   |                                 |
|     | 16.1   | Kontakt óra                                       | 28                              |
|     | 16.2   | Félévközi felkészülés órákra                      | 8                               |
|     | 16.3   | Felkészülés zárthelyire                           | 0                               |
|     | 16.4   | Zárthelyik megírása                               | 0                               |
|     | 16.5   | Házi feladat elkészítése                          | 14                              |
|     | 16.6   | Kijelölt írásos tananyag elsajátítása (beszámoló) | 0                               |
|     | 16.7   | Egyéb elfoglaltság                                | 0                               |
|     | 16.8   | Vizsgafelkészülés                                 | 10                              |
|     | 16.9   | <b>Összesen</b>                                   | <b>60</b>                       |
| 17. | Ellenőrző adat   |   | <b>Kredit * 30</b><br><b>60</b> |

|  |                          |                       |   |
|--|--------------------------|-----------------------|---|
| <b>A tantárgy tematikáját kidolgozta</b> |                          |                       |   |
| 18.                                      | Név                      | beosztás              | Munkahely (tanszék, kutatóintézet stb.) |
|  | <b>Dr. Kornai András</b> | <b>egyetemi tanár</b> | <b>Algebra Tanszék</b>                  |
|  |                          |                       |   |
|  |                          |                       |   |

|                        |                        |          |
|------------------------|------------------------|----------|
| <b>A tanszékvezető</b> |                        |          |
| 19.                    | Neve                   | aláírása |
|                        | <b>Dr. Nagy Attila</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.