

## BSc Matematika Alapszak, 2017.

Matematikai Intézet,  
Természettudományi Kar,  
Eötvös Loránd Tudományegyetem.

# Adatbányászat

- **Óraszám** ( $ea+gy$ ):  $2 + 2$
- **Specializáció**: elemző
- **Kredit** ( $ea+gy$ ):  $3 + 2$
- **Számonkérés**: kollokvium + gyak. jegy
- **Tárgykód** ( $ea, gy$ ): adatb\_1e0\_m17ea, adatb\_1e0\_m17ga
- **Ajánlott félév**: 6
- **Státusz**: kötelező

## Tantárgyfelelős

- Grolmusz Vince, Számítógéptudományi Tanszék, Matematikai Intézet.

## Előfeltételek

### *A gyakorlat előfeltételei:*

- **Erős**: Algebra2E (algebr2\*0\_m17ea)
- **Erős**: Véges matematika2E (vegmat2\*0\_m17ea)
- **Erős**: Gráfok és algoritmusok elméleteG-e (grafal1e0\_m17ga)

### *Az előadás előfeltételei:*

- **Gyenge**: a gyakorlat

## Megjegyzések

- **Pótlási lehetőség**: A félév végén, indokolt esetben, a gyakorlatvezető döntése alapján egy javító zárthelyi dolgozat írására van lehetőség.

### **A tematikát kidolgozta:**

- Grolmusz Vince, Számítógéptudományi Tanszék, Matematikai Intézet.

## Szükséges előismeretek

Lineáris algebra, elemi számelmélet, elemi gráfelmélet.

## A tantárgy célkitűzése

A tárgy célja az adatbányászat modern alapjainak felépítése

## Irodalom

- **Han, Kamber**: *Adatbányászat*.

## Tematika

- Adatok előfeldolgozása, transzponálása, tisztítása, redukálása, tömörítése, zajsűrése, dimenziócsökkentése. A dimenziócsökkentés módszerei és összehasonlításuk. Adattárházak, kezelésük. Az OLAP technológia. Az asszociációs szabályok bányászata. Gyakori halmazok tulajdonságai. Az apriori algoritmus és változatai. Hierarchikus asszociációs szabályok.
- Gyakori szekvenciák kinyerése. Klaszterezés: k-means és k-medoid. A CLARANS, BIRCH, DBSCAN algoritmusok és összehasonlításuk. Többdimenziós klaszterezés.