

Duomenų tyrimo egzamino temos (2023/24)

1. **Pagrindinės tikimybių teorijos ir informacijos teorijos sąvokos**
 - 1.1. Tikimybės. Atsitiktiniai dydžiai ir jų skirstiniai
 - 1.2. Įvykių sistemos
 - 1.3. Informacija ir entropija
 - 1.4. Atsitiktinių dydžių entropijos
2. **Pagrindiniai duomenų tyrimo uždaviniai ir sąvokos**
3. **Duomenys ir jų atributai**
 - 3.1. Pradinė duomenų analizė ir jų transformacijos
 - 3.2. Trūkstamosios reikšmės ir išskirtys
 - 3.3. Objektų artumo matai
4. **Kontroliuojamo mokymo uždaviniai: klasifikavimas**
 - 4.1. Sprendimų medžiai
 - 4.2. Modelio klaidos įverčiai
 - 4.3. Modelio perteklumas ir nepakankamumas
 - 4.4. Klasifikavimo taisyklės
 - 4.5. Artimiausių kaimynų metodas
 - 4.6. Bajeso klasifikatoriai
 - 4.6.1. Naivusis Bajeso klasifikatorius
 - 4.6.2. Bajeso tinklai
5. **Kontroliuojamo mokymo uždaviniai: skaitinė prognozė**
 - 5.1. Paprastosios tiesinės regresijos modelis
 - 5.2. Daugialypė tiesinė regresija
6. **Nekontroliuojamo mokymo uždaviniai**
 - 6.1. Asociacijos taisyklės
 - 6.2. Pirkėjo krepšelio uždavinys
 - 6.3. Klasterinė analizė
 - 6.3.1. Jungimo metodai
 - 6.3.2. K-vidurkių metodas