

## Veri Madenciliği Proje-2

Veri kümesindeki sekiz parametreye göre hastaların diyabetli olup olmadıklarını gösteren iki model tasarlayıp, bu modellerin sınıflama başarılarını karşılaştırmanız istenmektedir. İlgili veri kümesinde:

- A –H sütunu değerleri bağımsız parametrelerdir.
- I sütunu sınıf bilgisini tutmaktadır.
- A ve I sütunları hariç, diğer parametrelerdeki 0 (sıfır) değeri, bilinmeyen (eksik) değer olarak kabul edilmelidir. Sınıflama için K En Yakın Komşu, Bayes Ağları, Karar Ağaçları, Yapay Sinir Ağı algoritmalarından ikisini tercih edebilirsiniz.
- Karar ağacı için ID3 –CART –C4.5 –Random Forest vb yapılardan tercih ettiğinizi kullanabilirsiniz.
- KNN için en uygun K sayısını bularak modelinizi kurmanız gerekmektedir.
- İlgili modelleri kendiniz kodlayabilir veya WEKA – RapidMiner vb. hazır araçlar kullanabilirsiniz.