

## PROJE TESLİM TARİHİ: 07.04.2017

### Proje: Multi Thread Kullanarak Karar Ağacı Oluşturma

- **Proje masaüstü uygulaması olarak geliştirilecektir. Nesneye yönelik programlama ile geliştirmeniz istenmektedir. Kullanılabilecek diller: C++, C#, Python ve Java.**
- **Proje kaynak kodları (source code) CD içerisinde raporla beraber teslim edilecektir.**
- **Proje ile alakalı sorularınız için fidan.kaya@kocaeli.edu.tr ve ekin.ekinci@kocaeli.edu.tr adreslerine ulaşabilirsiniz.**

#### Proje Tanımı ve Yapmanız Gerekenler

Projede multi thread yapısını kullanarak size verilen veri seti üzerinden karar ağacı oluşturmanız istenmektedir. Karar ağacı oluşturma aşamasında C4.5 algoritması kullanılacaktır. Multi thread yapısını kullanmanız gereken iki ayrı durum bulunmaktadır. Bunlar;

- En iyi bölümlmeyi elde etmek için sadece sayısal veriler üzerinde thread kullanarak işlem yapmanız istenmektedir. Elde edilecek en iyi bölümleme durumu size alt seviyede oluşacak düğüm sayısını verecektir. En iyi bölümleme işlemi aşağıda detaylı olarak açıklanmıştır.
- Karar ağacında kök düğümünden sonraki tüm seviyelerde oluşacak düğüm sayısı kadar thread oluşturmanız ve her düğüm için bu threadleri eş zamanlı olarak çalıştırmanız istenmektedir.

**En İyi Bölümleme:** Veri setinde 3 özellik bulunmaktadır. Bu özelliklerden birincisi 30 ile 83 aralığında değer almaktadır; ikincisi 58 ile 69 aralığında değer almaktadır; üçüncüsü 0 ile 52 aralığında değer almaktadır. Bölümleme işlemi ilk özellik 50-60-70 için, ikincisi 62-63-64 için ve üçüncüsü 5-10-19 için bölümlenecektir. Örnek olarak ilk özellikte 50 değeri için yapılacak bölümleme  $<50$  ve  $\geq 50$  olarak yapılacaktır.

**Karar Ağacındaki Düğüm Sayısı:** Yani sizden binary tree oluşturmanız istenmektedir. Daha sonra bu iki dal için veri setinde bulunan diğer bir özelliği ağaca eklemek için oluşan iki düğüme ait iki thread çalıştırılacaktır. Veri setinde yer alan tüm özellikler ağaca yerleştirilene kadar bu işlem devam edecektir. Ağacın her seviyesinde yer alacak özelliği seçerken yine entropi hesabı yapmanız gerekmektedir.

**NOT: Kullanılacak veri seti “haberman.txt” dir. Veri setinin açıklamasını içeren “haberman\_aciklama.txt” dosyası proje klasöründe verilmiştir.**

### Proje Değerlendirme Yönergesi Açıklamalar

- **PROJE RAPORU:** Her proje sonunda teslim edilmesi gereken projenizi her yönüyle açıklayacağınız bir dokümandır.
  - **Problem Tanımı:** Proje kapsamında sizden çözüm bulmanız beklenen problem ile ilgili açıklama yapılması gerekmektedir. Burada amaç projenin ne kadar anlaşıldığını test etmektir.
  - **Yapılan Araştırmalar:** Proje geliştirilmesi aşamasında karşılaşılan sorunlara nasıl çözümler bulunduğu ve bu konularda yapılan araştırmalar açıklanmalıdır.
  - **Tasarım**
    - **Akış şeması:** Proje içerisinde yer alan algoritma ve işlemleri şekilsel

- olarak ifade edecek şema oluşturmanız beklenmektedir.
- **Yazılım mimarisi:** Projenin kodlanması aşamasında kullanılacak kod yapısı ve geliştirme aşamalarını gösteren bir yapı hazırlanması beklenmektedir.
- **Veri tabanı diyagramı:** Projeye ait ER diyagramının oluşturulması beklenmektedir.
- **Genel Yapı:** Projenizi genel yapısı bakımından her yönüyle özetlemeniz gerekmektedir.
- **Referanslar:** Proje geliştirilirken ve araştırma aşamasında faydalanılan kaynaklar rapor dokümanının en altında listelenmeli ve doküman içerisinde de ilgili yerlerde indekslenmelidir.

**Referans formatı aşağıda verilen örneklerle uygun olmalıdır.**

**Kitap, çok yazarlı**

Larson, G. W., Ellis, D. C., & Rivers, P. C. (1984). Essentials of chemical dependency counseling. New York: Columbia University Press.

**Report from a private organization (author & publisher same)**

National League for Nursing. (1990). Self-study report for community health organizations (Pub. No. 21-2329). New York: Author.

**Unpublished master's thesis**

Paulosky, K. A. (1997). Knowledge and attitudes of pain and activities of nurse administrators. Unpublished master's thesis, Northern Michigan University, Marquette, Michigan.

**Article in a journal (continuous pagination throughout volume)**

Burke, R. J., Shearer, D., & Deszca, E. (1984). Correlates of burnout phases among police officers. Group and Organizational Studies, 9, 451-466.

**Article in a Popular Magazine**

Caloyianis, N. (1998, September). Greenland sharks. National Geographic, 194, 60- 71.

**Web Site**

[http://en.wikipedia.org/wiki/Neural\\_network](http://en.wikipedia.org/wiki/Neural_network) (Access date: 07.10.2013)

**İNTİHAL: İNTERNETTEN ALINAN KOD PARÇACIKLARI MUTLAKA KOD İÇERİSİNDE BELİRTİLECEK VE AÇIKLAMA SATIRI İLE KAYNAK GÖSTERİLECEKTİR. AKSİ DURUMDA KOPYA OLARAK DEĞERLENDİRİLECEKTİR. KOPYA ÇEKTİĞİ YA DA KOPYA VERDİĞİ TESPİT EDİLEN ÖĞRENCİLER SUNUMA ALINMAYACAKTIR.**

## **PROJELER İKİ KİŞİLİK GRUPLAR HALİNDE YAPILACAKTIR!**

**Proje gruplarının oluşturulacağı link:**

**<https://docs.google.com/spreadsheets/d/1yNKVwf7N6p6dA-EyJILQqRQcTBzqrSSqeQ1ABHXIsmQ/edit?usp=sharing>**