****

**COMP435 – Makine Öğrenmesi ödev raporu**

**Makine Öğrenimi Algoritmalarını Kullanarak Orman Yangını Oluşumunu ve Büyüklüğünü Tahmin Etmek**

**Serhat Erdoğan 190301026**

**Çalışmanın amacı:** Makine Öğrenimi Algoritmalarını Kullanarak Orman Yangını Oluşumunu ve Büyüklüğünü Tahmin Etmek.

Orman yangınlarıyla mücadele etmek için yangın çıkana kadar beklemek yerine orman yangınlarını azaltmaya odaklanmaya ihtiyaç vardır. Bu araştırma, dünyadaki orman yangınlarının azalmasına yol açacak doğru tahmin yoluyla dünya ormanlarının korunmasına büyük katkı sağlayacaktır. Bu çalışmanın amacı, mevcut verileri kullanarak doğal ortamda ortaya çıkan yangınları tahmin etmek ve yangının büyüklüğünü tahmin etmektir. Bu çalışmada, California Eyaleti'nde meydana gelen yangınların özelliklerini inceleyerek, yangının ne zaman ve nerede ortaya çıkacağını tahmin etmek amaçlanmıştır.

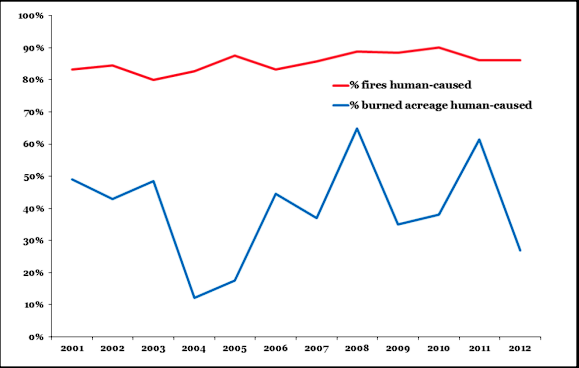
Bu çalışmada, aşağıdaki yöntemler kullanılmıştır:

Yangın verileri toplanmış ve ön işleme tabi tutulmuştur.

makine öğrenimi yöntemi kullanılarak model oluşturulmuştur.

Model, eğitim ve test verileri kullanılarak değerlendirilmiştir.

Modelin performansı değerlendirilmiş ve sonuçlar yorumlanmıştır.

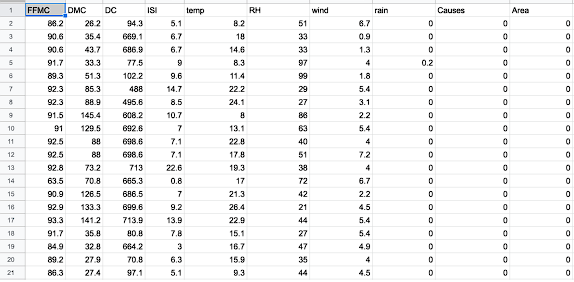


**2001-2012 yılları arasında Amerika Birleşik Devletleri'ndeki Orman Yangınlarının nedenleri**

Bu çalışmada, yangının ne zaman ve nerede ortaya çıkacağını tahmin etmek için Yapay sinir ağları makine öğrenimi algoritması kullanılmıştır.

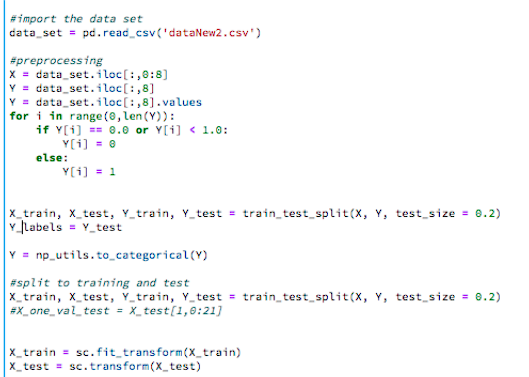
**Yapay sinir ağları nedir ?**

Yapay sinir ağları (YSA), insan beyninin bilgi işleme tekniğinden esinlenerek geliştirilmiş bir bilgi işlem teknolojisidir. YSA ile basit biyolojik sinir sisteminin çalışma şekli taklit edilir. Yani biyolojik nöron hücrelerinin ve bu hücrelerin birbirleri ile arasında kurduğu sinaptik bağın dijital olarak modellenmesidir. Nöronlar çeşitli şekillerde birbirlerine bağlanarak ağlar oluştururlar. Bu ağlar öğrenme, hafızaya alma ve veriler arasındaki ilişkiyi ortaya çıkarma kapasitesine sahiptirler. Diğer bir ifadeyle, YSA'lar, normalde bir insanın düşünme ve gözlemlemeye yönelik doğal yeteneklerini gerektiren problemlere çözüm üretmektedir. Bir insanın, düşünme ve gözlemleme yeteneklerini gerektiren problemlere yönelik çözümler üretebilmesinin temel sebebi ise insan beyninin ve dolayısıyla insanın sahip olduğu yaşayarak veya deneyerek öğrenme yeteneğidir.



Yukarıdaki tablo modelimiz için sıcaklık, bağıl nem, nem içeriği, rüzgar hızı ve yağmur miktarı gibi bilgiler içeriyor. Modelimizi eğitmek için bu verisetini kullanacağız.

**Modelin eğitilmesi**

****

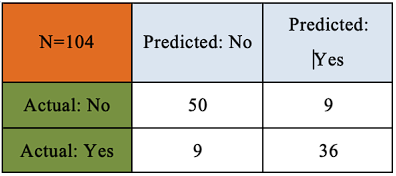
kodda gösterildiği gibi scikit-learn kitaplığının fit\_transform yöntemi kullanılarak eğitimden önce optimize edildi. Daha sonra, özellikler scikit'in train\_test\_split yöntemi kullanılarak iki kümeye ayrıldı Özelliklerin yüzde 80'i modeli eğitmek için kullanılırken, kalan yüzde 20'lik kısım ise modelin performansını test etmek için kullanıldı.

**Modelin test edilmesi**



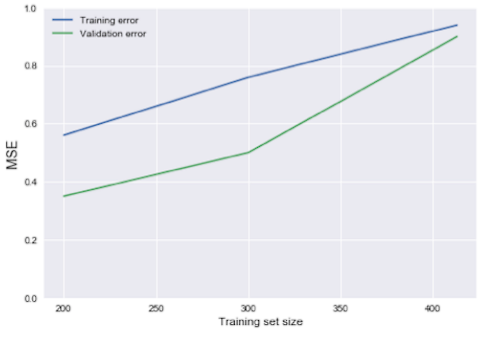
Karışıklık matrisi, eğitimli modeli test etmek için toplam 104 özelliğin kullanıldığını gösterir. Tahminlerin gerçek etiketlerle karşılaştırılması, 59 Hayır etiketinden 50'sinin gerçekte Hayır ve 9'unun Evet olarak tahmin edildiğini ve 45 Evet etiketinden, 36'sı aslında Evet olarak tahmin edildi ve 9'u Hayır olarak tahmin edildi. Bu tahmin, yaklaşık 0,83 model doğruluk puanı ve 0,83 ortalama kesinlik anlamına gelir.

Gerçek pozitif, gerçek negatif, yanlış pozitif ve yanlış negatif değerleri, aşağıdaki tabloda gösterildiği gibi karışıklık matrisinden belirlendi.

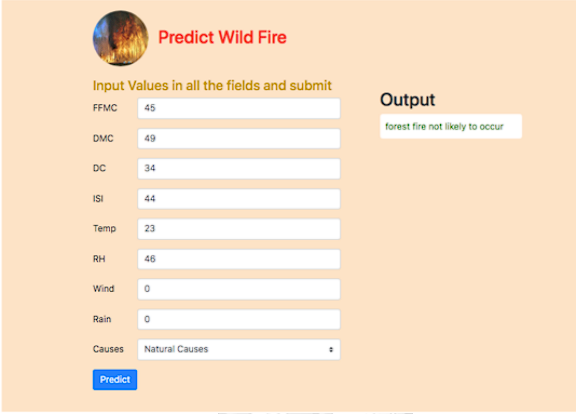


Tahmin Doğruluğu yüzde 82,69 olarak belirlenmiştir.

Aşşağıda gösterildiği gibi sınıflandırıcının performansını görselleştirmek için kullanılan farklı boyutlardaki özelliklere sahip Yapay sinir ağı modeli için bir Öğrenme eğrisi çizilmiştir. Model daha fazla özellik kullanıldıkça performansı arttı. Bu, gelecekte özelliklerin arttırılmasının, eğitim eğrisi ile doğrulama eğrisinin yakınsamasını sağlayacağı anlamına gelir.



Kullanıcıya, modeli sorgulamak için tahmin parametrelerini girmesi için bir web arayüzü sağlanır. Model daha sonra, tahmin parametrelerini girmek için kullanılan aynı web arabirimi aracılığıyla görüntülenen sonucu tahmin etmek için yüklenir. Aşşağıdaki gibi giriş/çıkış kullanıcı arayüzünü gösterir.



**SONUÇ**

Veriler, hem meteorolojik kaynakları hem de 2000 ile 2020 arasındaki içeriğe odaklanan orman yangını veritabanlarını içeren güvenilir klimatolojik kaynaklardan toplanmıştır; Orman yangınlarının tutuşmasına katkıda bulunan tüm ana bileşenler dikkate alındı. Veriler önceden işlendi ve eğitim ve test setlerine ayrıldı. Eğitim seti, Yapay sinir ağları modelini eğitmek için kullanıldı. Test seti, yüzde 82,69 doğruluk sağlayan modeli test etmek için kullanıldı.

Bu çalışma, Yapay sinir ağlarının orman yangınlarını tahmin etmek için kullanılabileceğini ve orman yangınlarının önlenmesi ve kurtarılmasında yardımcı olacağını göstermektedir. Toplam 20.000 veri hacmi toplandı ancak modeli eğitmek için yalnızca 414 (%80) kayıt kullanıldı ve kalan 103 (%20) kayıt test etmek ve doğrulamak için kullanıldı.