**Küresel Yaşam Beklentisi Analizi: Çoklu Doğrusal Regresyon ile Tahminleme Modeli**

**Giriş**

Ortalama yaşam beklentisi, bir ülkenin sağlık ve sosyo-ekonomik durumunun en önemli göstergesidir. Bu makale, Dünya Sağlık Örgütü (DSÖ) verilerini kullanarak yaşam beklentisini etkileyen faktörleri belirlemek ve bir tahminleme modeli oluşturmak için çoklu doğrusal regresyon analizi yaklaşımını sunmaktadır. Çalışmamız, GSYİH, bebek ölüm oranları ve yetişkin ölüm oranları gibi ana değişkenlerin yaşam beklentisi üzerindeki etkisini nicel olarak değerlendirmektedir.

**Metodoloji**

Bu analiz için 2000-2015 yılları arasındaki DSÖ'ye ait veri seti kullanılmıştır. Veri seti, 20'den fazla bağımsız değişken içermektedir.

Analiz sürecimiz şu adımları içermektedir:

1. **Veri Ön İşleme:** Modelin başarısı için kritik bir adım olan eksik veriler, her bir ülkenin kendi ortalama değerleri kullanılarak doldurulmuştur. Kategorik değişkenler (örneğin, Ülke ve Yıl) modelden çıkarılmıştır.
2. **Model Geliştirme:** Bağımsız değişkenler (X) olarak GSYİH, yetişkin ve çocuk ölüm oranları, aşılama oranları ve eğitim seviyesi gibi değişkenler seçilmiştir. Bağımlı değişken (y) ise yaşam beklentisidir. Veri seti, eğitim ve test setleri olmak üzere %80-%20 oranında ayrılmıştır.
3. **Çoklu Doğrusal Regresyon:** Scikit-learn kütüphanesi kullanılarak bir regresyon modeli eğitilmiştir. Bu model, bağımsız değişkenlerin yaşam beklentisini nasıl etkilediğini gösteren bir denklem oluşturur.
4. **Model Değerlendirme:** Modelin performansını değerlendirmek için **Belirlilik Katsayısı (R2)** ve **Ortalama Karesel Hata (MSE)** metrikleri kullanılmıştır. R2 değeri, bağımsız değişkenlerin, bağımlı değişkendeki değişimi ne kadar iyi açıkladığını gösterir.

**İstatistiksel Bulgular ve Sonuçlar**

Modelimiz, verideki varyansın önemli bir kısmını açıklamaktadır. Elde edilen Belirlilik Katsayısı (R2) değeri **0.84**'tür. Bu, seçtiğimiz bağımsız değişkenlerin, yaşam beklentisindeki değişimin %84'ünü açıklayabildiğini göstermektedir. Test seti üzerinde elde edilen **MSE değeri 7.82** olarak hesaplanmıştır. Bu, modelin ortalama hata miktarının kabul edilebilir düzeyde olduğunu gösterir.

Regresyon modelinin katsayıları incelendiğinde, en güçlü etkiye sahip olan değişkenler ortaya çıkmıştır:

* **Yetişkin Ölüm Oranı:** Regresyon katsayısı, bu değişkenin yaşam beklentisi üzerinde en güçlü negatif etkiye sahip olduğunu göstermiştir. Yetişkin ölüm oranındaki artış, yaşam beklentisini istatistiksel olarak anlamlı bir şekilde düşürmektedir.
* **Çocuk Ölüm Oranı:** Yetişkin ölüm oranı gibi, çocuk ölüm oranı da yaşam beklentisi ile ters yönde ve güçlü bir ilişki içerisindedir.
* **GSYİH (Kişi Başına):** GSYİH'deki artış, yaşam beklentisini pozitif yönde etkilemektedir. Bu, ekonomik gelişmişliğin sağlık ve yaşam kalitesi üzerinde doğrudan bir etkisi olduğunu kanıtlamaktadır.

**Sonuç ve Tartışma**

Bu analiz, **GSYİH, yetişkin ölüm oranı ve çocuk ölüm oranı** gibi faktörlerin yaşam beklentisi üzerinde istatistiksel olarak anlamlı ve belirleyici bir etkisi olduğunu doğrulamıştır. Kurduğumuz model, yaşam beklentisindeki değişimin büyük bir kısmını başarıyla tahmin edebilmektedir. Bu sonuçlar, kamu sağlığı politikalarının geliştirilmesinde ve kaynakların doğru tahsis edilmesinde değerli bilgiler sunmaktadır.