

İnme Riskinin Tahmini: Sağlık Verileri ve Makine Öğrenimi

Mahmut Sami Yılmaz
Seda Yeler



Giriş

- Bu projede, sağlık verileri kullanarak inme riskini tahmin etmek amacıyla bir makine öğrenimi modeli geliştirildi. Erken teşhis ve önleyici tedbirler hayati öneme sahiptir. Bu sunumda, proje adımları ve elde edilen sonuçlar ele alınacaktır.



Veri ve Yöntemler

- Veri Ön İşleme: Eksik değerlerin doldurulması, kategorik değişkenlerin sayısal formata dönüştürülmesi.
- Veri Dengeleme: SMOTE kullanarak sınıf dengesizliğinin giderilmesi.
- Özellik Ölçekleme: StandardScaler ile standartlaştırma.
- Modelleme: Logistic Regression ile model eğitimi.
- Değerlendirme: Doğruluk, hassasiyet, duyarlılık ve F1 skoru ile performans değerlendirme.

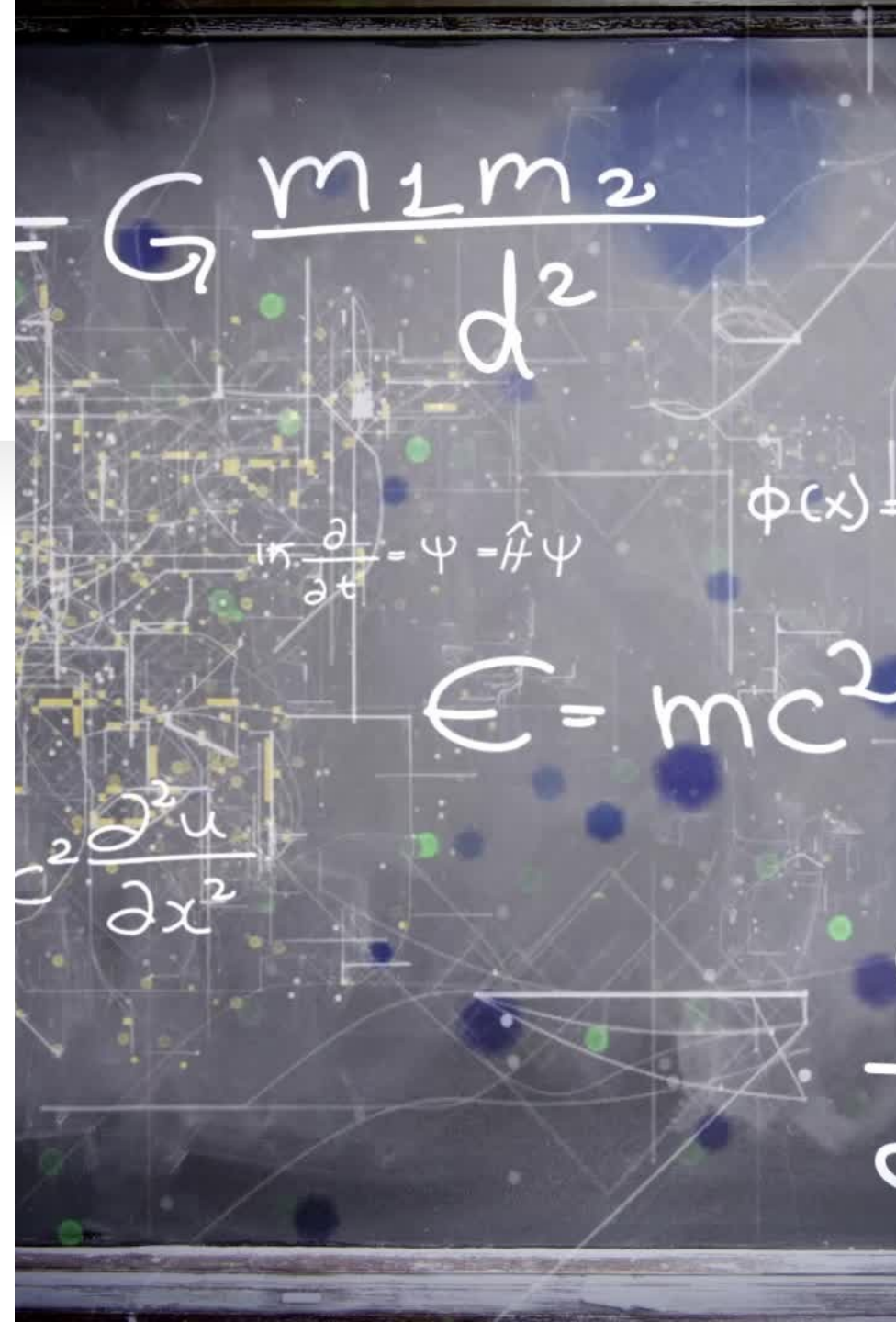
Veri Seti Hakkında

- Veri seti, inme geçiren ve geçirmeyen bireylerin çeşitli sağlık ve demografik bilgilerini içermektedir. Değişkenler arasında cinsiyet, yaş, yüksek tansiyon, kalp hastalığı, evlilik durumu, çalışma durumu, yaşam yeri, ortalama glukoz seviyesi, BMI, sigara içme durumu ve inme durumu bulunmaktadır.



Model Performansı

- Modelin performansı şu metriklerle değerlendirilmiştir:
- Doğruluk: 0.9568
- Hassasiyet: 0.9944
- Duyarlılık: 0.9186
- F1 Skoru: 0.9550
- Bu metrikler, modelin inme tahminindeki başarısını göstermektedir.



Tartışma

- Bu çalışma, Logistic Regression modelinin veri seti üzerinde oldukça başarılı bir şekilde çalıştığını göstermektedir. Model, inme tahmininde yüksek bir doğruluk ve F1 skoru ile dikkat çekmektedir. Ancak, daha karmaşık modellerin kullanılması ve veri setine daha fazla özellik eklenmesi, tahmin performansını daha da artırabilir.



Sonuç ve Gelecek Çalışmalar

- Sonuç olarak, Logistic Regression modeli ile inme riski yüksek doğrulukla tahmin edilmiştir. Gelecekte, daha karmaşık modellerin denenmesi ve veri setine yeni özelliklerin eklenmesiyle daha yüksek doğruluk elde edilebilir.

Soru ve Cevap

- - Sorularınız var mı?