

NFM VE FiBiNET MODELLERİNİN BİRLEŐTİRİLMESİ

Aziz Sarıtaő

NFM VE FIBINET MODELLERİNİN BİRLEŞTİRİLMESİ

Giriş

CTR (Click-Through Rate) tahmini, çevrimiçi reklamcılık ve öneri sistemlerinde önemli bir sorundur. Bu, bir kullanıcının belirli bir reklama veya öneriye tıklama olasılığını tahmin etmeyi içerir. Temel amaç, kullanıcı etkileşimini ve geliri en üst düzeye çıkarmak için kullanıcılara daha olası tıklamalar alabilecek reklamlar veya içerikler sunmaktır. Çevrim- içi reklam bağlamında CTR tahmini, reklamverenlere ve yayıncılara reklam yerleşimlerini optimize etme ve kaynakları etkili bir şekilde tahsis etme konusunda yardımcı olur. Bir kullanıcının reklama tıklama olasılığını doğru bir şekilde tahmin ederek, reklamverenler reklam yerleşimleri için daha stratejik tekliflerde bulunabilirken, yayıncılar daha ilgili içerikleri görüntüleyerek kullanıcı deneyimini artırabilir. CTR tahmini problemi doğasında zorlu bir yapıya sahiptir oldukça büyük verisetlerinde çalışmayı gerektirebilir. Ayrıca, model eğitimi için kullanılan veri genellikle dengesizdir çoğu durum tıklamama durumlarını içerir. Bu sorunu ele almak için genellikle lojistik regresyon, karar ağaçları ve daha gelişmiş modeller olarak gradient boosting ve sinir ağları gibi makine öğrenimi teknikleri kullanılır. Genel olarak, CTR tahmini çevrimiçi reklamcılığın ve öneri sistemlerinin optimize edilmesinde önemli bir rol oynar, dijital pazarlama stratejilerinin etkinliğine ve kullanıcılara kişiselleştirilmiş içerik sunumuna katkıda bulunur.

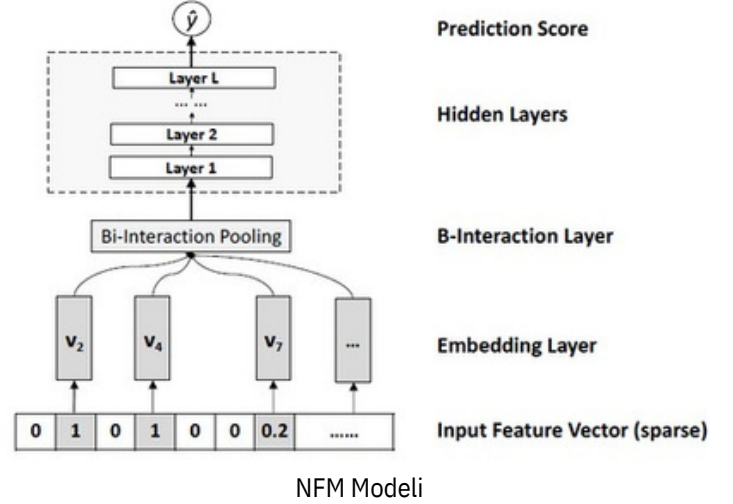
1 Metodoloji

1.1 NFM Modeli

NFM, veya Neural Factorization Machine, işbirliğine dayalı filtreleme (collaborative filtering) ve öneri görevleri için tasarlanmış bir makine öğrenimi modelidir. NFM, faktörleştirme makinelerinin ve sinir ağlarının avantajlarını birleştirerek verideki hem lineer hem de lineer olmayan ilişkileri yakalamayı amaçlar. NFM Modelinde kullanılan katmanlar:

Faktörleştirme Makineleri (FM): FM modelleri, özellikler arasındaki etkileşimleri lineer bir şekilde yakalamak için etkilidir. Her çift için ağırlık atayarak özellikler arasındaki ikili etkileşimleri modellemektedir. Bu, özellikler arasındaki ilişkileri özellikle seyrek veri setlerinde yakalamak için kullanışlıdır.

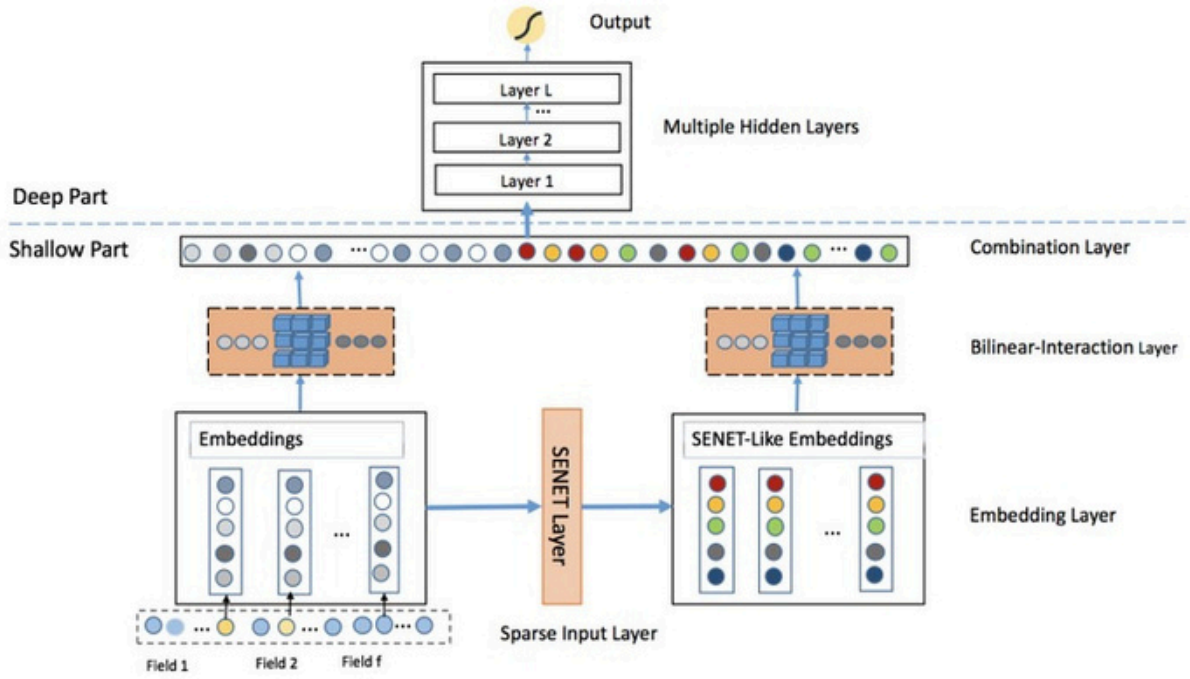
Sinir Ağları: NFM, sinir ağlarını kullanarak verideki lineer olmayan kalıpları ve karmaşık ilişkileri yakalamayı amaçlar. Sinir ağı bileşeni, giriş özelliklerinin yüksek düzeyde temsiliyi öğrenir ve yalnızca lineer etkileşimlerle etkili bir şekilde modellenemeyen karmaşık bağımlılıkları yakalamak için kullanılır.



FM ve Sinir Ağlarının Birleşimi: NFM, FM'nin etkileşimlerin lineer bileşenini modellemek ve sinir ağlarını lineer olmayan kalıpları yakalamak için birleştirmektedir. FM, etkileşimlerin lineer bileşenini modellemek için kullanılırken, sinir ağı, lineer olmayan desenleri yakalamak ve özellikler arasındaki karmaşık ilişkileri anlamak için kullanılır. Gömme Katmanları: NFM genellikle giriş özelliklerini sürekli vektörlere dönüştürmek için gömme katmanları kullanır. Bu gömme, eğitim süreci sırasında öğrenilen ve modelin farklı özellikler arasındaki içsel özellikleri ve ilişkileri anlamasına yardımcı olan özelliklerdir. Sonuç olarak, NFM, öneri görevleri için etkili bir çözüm sunan faktörleştirme makinelerinin ve sinir ağlarının gücünü birleştiren hibrit bir modeldir. Lineer ve lineer olmayan ilişkileri yakalama

kollanıldıkin yeteneği, özellikle karmaşık etkileşimlerini doğru bir şekilde anlamının önemli olduğu işbirliğine dayalı filtreleme senaryoları için uygundur. NFM'de diğer modellerden farklı olarak kullanılan Bi-Interaction pooling bileşenindedir. Bi-Interaction havuzlama kullanımı, düşük seviyede ikinci dereceden özellik etkileşimlerini yakalar, bu da birleştirme işleminden daha bilgilidir. Bu, NFM'nin ardından gelen gizli katmanların çok daha kolay bir şekilde yararlı yüksek dereceli özellik etkileşimlerini öğrenmesine büyük ölçüde katkıda bulunur.

1.2 FiBiNET Modeli

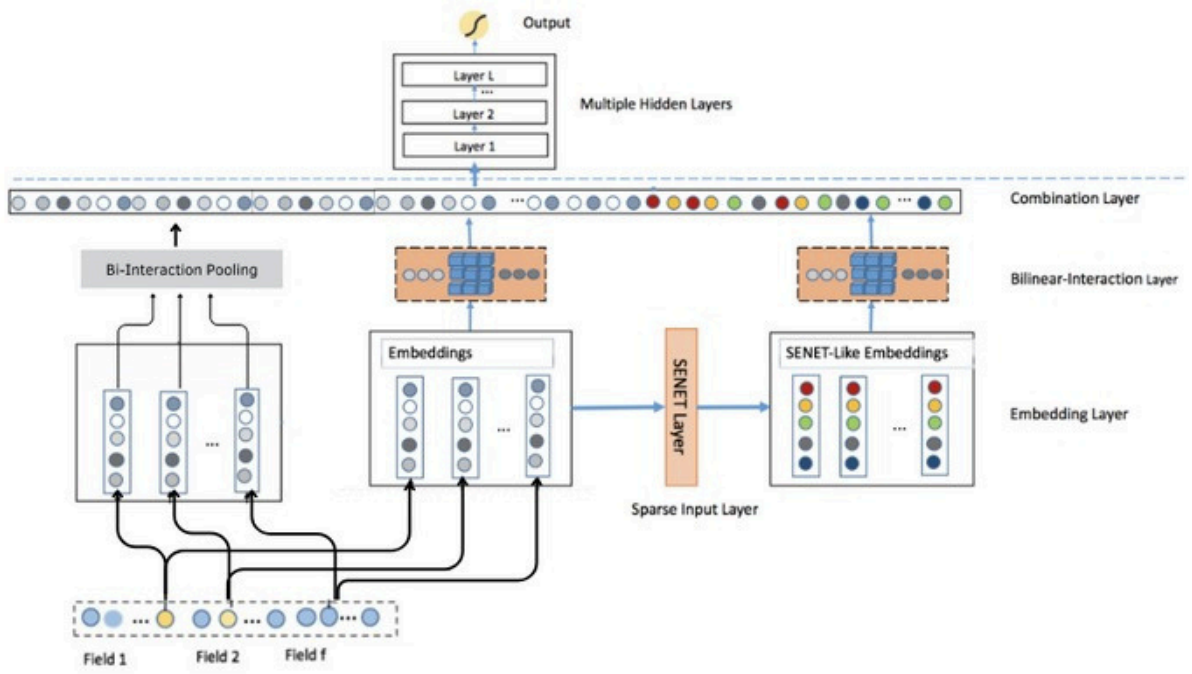


FiBiNET Modeli

FiBiNET, (Feature Importance and Bilinear Feature Interaction Network) özelliklerin önemini ve ince taneli özellik etkileşimlerini dinamik olarak öğrenmek üzere tasarlanmış bir modeldir. FiBiNET, özelliklerin önemini Squeeze-Excitation network (SENET) mekanizması ile öğrenir. Bu mekanizma, her özelliğin ağırlığını, özelliğin kendisi ve diğer özelliklerle olan ilişkisine göre ayarlar. Özellik etkileşimlerini ise ikili fonksiyon ile öğrenir. FiBiNET, her bir çift özelliğin etkileşimini bilerek bir bilinear özellik etkileşim katmanı kullanır. Bu katman, her bir özelliğin diğer tüm özelliklerle olan etkileşimini tek bir vektörde birleştirir. Bu özellik, FiBiNET'in model karmaşıklığını azaltmasına ve performansını artırmasına yardımcı olur. Bu fonksiyon her iki özelliğin de değerini dikkate alarak, özellikler arasındaki çapraz etkileri yakalar. Bu sayede FiBiNET modeli gerçek veri kümesinde yapılan deneylerde diğer yüzeysel modellerden daha iyi performans göstermiştir. FiBiNET, DeepFM ve extreme deep factorization machine (XdeepFM) gibi diğer son teknoloji derin modellerden de tutarlı bir şekilde daha iyi performans gösterir. Özellik mühendisliği gerektirmeyen, ölçeklenebilir ve etkili bir modeldir. FiBiNET modeli, özellikle büyük ve karmaşık veri setleri üzerinde etkili bir şekilde çalışabilen, özellik etkileşimlerini yakalamak konusunda güçlü bir yetenek sunan bir derin öğrenme modelidir.

1.3 NFM Modeli ve FiBiNET Modelinin Birleştirilmesi

FibiNet Modelinde bulunan Bilinear Interaction ve SENET katmanlarının NFM modeline eklenmesi ile oluşturulan yeni NFM modelinde başarı oranı normal NFM modeline göre daha yüksek olmuştur. NFM Modeli özellik etkileşimlerini öğrenmek için bir faktörleştirme makinesi ile bir çok katmanlı algılayıcıyı birleştirir. NFM Modeli, özelliklerin ikili etkileşimlerini yakalamak için gizli vektörlerin çarpımını kullanır. Bu model, özellik etkileşimlerinin doğrusal olmadığını varsayar ve bu nedenle çok katmanlı algılayıcıyı kullanarak daha karmaşık etkileşimleri öğrenebilir. FiBiNET Modeli özelliklerin önemini ve ince taneli özellik etkileşimlerini dinamik olarak öğrenmek için tasarlanmıştır. FiBiNET Modeli, özelliklerin önemini Squeeze-Excitation network (SENET) mekanizması ile öğrenebilir. Özellik etkileşimlerini ise ikili fonksiyon ile etkili bir şekilde öğrenebilir. NFM Modelinin kullandığı Inner Product Interaction katmanının ürettiği çıktılar ile FiBiNET modelinin Bilinear Interaction Katmanının ürettiği çıktılar birleştirilerek DNN katmanına girilir. Bu sayede çok katmanlı DNN hesaplamaları sonucu daha yüksek bir AUC elde edilir.



**NFM+FiBiNET Model
Diyagramı**

2 Sonuçlar

2.1 Ablation Study

Durumlara Göre AUC ve Logloss Değerleri

NFM Modeli için:

NFM Modeli Normal hali ile çalıştırıldığında AUC: 0.9807
logloss:0.1988 olarak ölçüldü.

Modelimizde bi_pooling_layer (InnerProductInteraction) katmanı olmadan çalıştırıldığında AUC: 0.9739 logloss:0.1967 olarak ölçüldü.

lr_layer (Logistic Regression) katmanı olmadan çalıştırıldığında elde edilen sonuçlar: AUC: 0.9783 logloss:0.1792

MLP_Block (DNN) katmanı olmadan çalıştırıldığında elde edilen sonuçlar ise : AUC: 0.5057 logloss:0.6711

Bu sonuçlara göre bütün katmanların AUC değerine etki ettiği görülmektedir.

Özellikle MLP_Block kullanılmadığında AUC değeri oldukça fazla düşmektedir.

Durum	AUC	Logloss
lr_layer olmadan	0.97837	0.17922
bi_pooling_layer olmadan	0.97390	0.19677
MLP_Block olmadan	0.50573	0.67113
sadece MLP_Block + lr_layer	0.98041	0.13528
sadece bi_pooling_layer+lr_layer	0.81468	0.58320

2.1 Birleştirilmiş Model Sonuçlarının Karşılaştırılması

NFM Modeli Normal hali ile çalıştırıldığında AUC: 0.9807
logloss:0.1988 olarak ölçüldü.
Aynı Parametrelerde sadece FiBiNET çalıştırılınca: AUC: 0.98055
logloss:0.19845 olarak ölçüldü.
Oluşturulan NFM+FiBiNET modeli çalıştırıldığında AUC: 0.982136
logloss:0.170597 olarak ölçüldü.
Tüm durumlarda Kullanılan Parametreler:
model: NFM
loss: 'binary_crossentropy'
metrics: ['logloss', 'AUC']
task: binary_classification
optimizer: adam
learning_rate: 0.001

embedding_regularizer: 0.01
net_regularizer: 0
batch_size: 512
embedding_dim: 10
hidden_units: [400, 400, 400]
hidden_activations: relu
net_dropout: 0.1
batch_norm: False
epochs: 100
shuffle: True
seed: 2023
monitor: 'AUC'
monitor_mode: 'max'

Durum	AUC	Logloss
NFM	0.98071	0.19886
FiBiNET	0.98055	0.19845
NFM + FiBiNET	0.98213	0.17059

3 Tartışma

Görüldüğü üzere modeller aynı parametrelerle frappe datasetinde test edildiğinde birleştirilmiş modelin sonuçları normal modellerden daha başarılı sonuç vermiştir.
HFM modelinin ablasyon sonuçlarını incelediğimizde ise modelin katmanlarına göre değişen AUC değerleri gözlemlenmiştir.
Özellikle DNN kullanıldığında AUC değeri oldukça artan model InnerProductInteraction ve LR katmanları sayesinde de doğruluğunu artırmaktadır.
Bu modele FiBiNET modelinde bulunan BiLinearInteraction ve SqueezeExcitation katmanlarını da eklediğimiz zaman modelimiz 0.982136 AUC ve 0.17059 Logloss değeri ile BARS Benchmark'taki Temel HFM modelin AUC değerini geçmektedir.