

# Şiir Türlerinin Transformatör Model Eğitimi İle Sınıflandırılması

Hilal Yüce  
Bilişim Sistemleri Mühendisliği  
Kocaeli Üniversitesi  
Kocaeli, Türkiye  
221307070

**Öz—**Bu raporda Yazılım Geliştirme Laboratuvarı-I dersinin II. Projesi kapsamında, Türkçe şiirleri sınıflandırmak amacıyla beş farklı Transformer tabanlı modelin geliştirilmesi ve değerlendirilmesi ele alınmıştır. Proje, Google Colab platformunda gerçekleştirildi ve Python programlama dili kullanılarak uygulanmıştır. Modeller, eğitim ve test veri setleri üzerinde doğruluk (Accuracy), geri çağırma (Recall), kesinlik (Precision), F-Score, duyarlılık (Sensitivity), özgüllük (Specificity) ve ROC eğrisi gibi performans metrikleri ile analiz edilmiştir. Sonuçlar, modellerin beş farklı şiir kategorisi için etkili bir şekilde sınıflandırma yapabileceğini göstermektedir. Kullanılan araçlar arasında Huggingface Transformers, PyTorch ve scikit-learn kütüphaneleri yer almaktadır.

**Anahtar Kelimeler—**Transformer, şiir, veri seti, model, sınıflandırma, eğitim.

## I. GİRİŞ

Doğal Dil İşleme, metin tabanlı verilerin analizi ve anlamlandırılması için önemli bir alan haline gelmiştir. Bu projede, Türkçe şiirlerin belirli kategorilere (epik, pastoral, lirik, satirik ve felsefi) sınıflandırılması hedeflenmiştir. GPT-2, BERT, RoBERTa, ELECTRA ve DistilBERT modelleri, güçlü dil işleme yetenekleri nedeniyle seçilmiş ve projenin temel modelleri olarak kullanılmıştır.

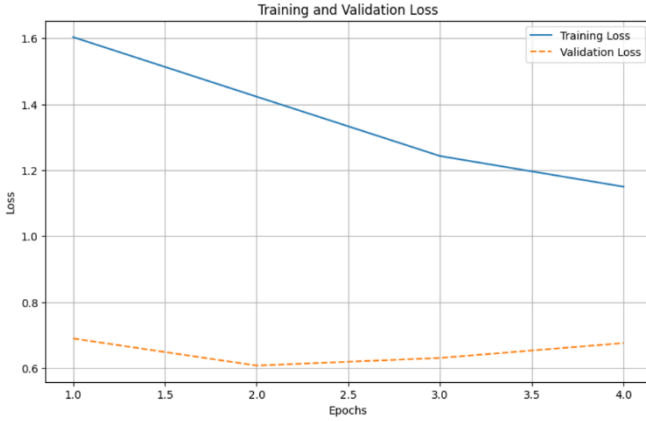
## II. MODELLERİN GELİŞTİRİLMESİ

Modellerin geliştirilmesi sürecinde ilk olarak, eğitim veri setleri dikkatlice hazırlanmış ve eksik ya da uyumsuz veriler kontrol edilerek temizlenmiştir. Veri setleri, eğitim ve test olmak üzere ikiye ayrılmıştır. Model hiperparametreleri (öğrenme oranı, batch boyutu, epoch sayısı) titizlikle optimize edilerek her bir modelin en iyi performansı göstermesi hedeflenmiştir. Ayrıca, tokenizasyon işlemleri sırasında her modelin doğal dil işleme yeteneklerine uygun veri formatları oluşturulmuş, bu sayede modellerin eğitim süreçleri verimli bir şekilde tamamlanmıştır. Projede, her bir şiir türü belirli bir sınıf numarası ile eşleştirilmiştir. Sınıflar ve bu sınıfların temsil ettiği şiir türleri aşağıda sıralanmıştır:

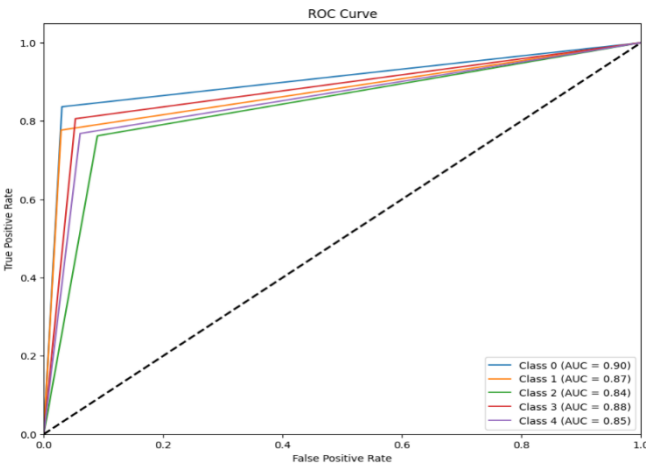
- Class 0:** Epik Şiir (Kahramanlık ve vatan sevgisi konulu şiirler)
- Class 1:** Pastoral Şiir (Doğa, kırsal yaşam ve tabiat güzelliklerini konu alan şiirler)
- Class 2:** Lirik Şiir (Duygusal ve bireysel hislerin ön planda olduğu şiirler)
- Class 3:** Satirik Şiir (Toplumsal veya bireylerin aksayan yönlerini eleştiren şiirler)
- Class 4:** Felsefi Şiir (Hayat, zaman, inanç ve ahlak gibi derin temalar içeren şiirler)

## A. BERT

Bert-base-turkish-cased modeli kullanılmıştır. Modelin sınıflandırma yapabilmesi için son katmanı 5 sınıfa ayırarak edebilecek şekilde uyarlanmıştır. Her epoch'ta modelin doğruluğu ve loss yakından izlenmiştir. Veri setleri, her bir satırda şiir metni ve bu metne karşılık gelen etiket bilgilerini içermektedir. Eğitim sonunda model, 4 epoch'ta optimum performansa ulaşmıştır.

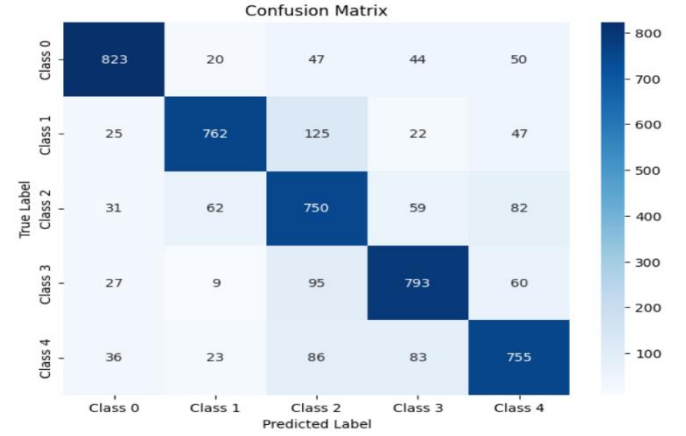


Bu grafik, modelin eğitim sürecinde elde edilen eğitim ve doğrulama kayıplarını göstermektedir. Mavi çizgi, modelin eğitim kaybını temsil ederken turuncu kesik çizgi doğrulama kaybını temsil etmektedir. Grafikte, eğitim kaybının her epoch sonunda azaldığı ve modelin öğrenme sürecinin devam ettiği gözlemlenmektedir. Ancak doğrulama kaybında küçük bir artış görülmektedir; bu, modelin aşırı öğrenme eğiliminde olabileceğini göstermektedir. Bu durumu önlemek için Early Stopping uygulanmıştır.



Bu grafik, modelin sınıf bazında ROC eğrilerini ve her sınıf için AUC değerlerini göstermektedir. ROC eğrisi, modelin yanlış pozitif oranına

karşılık doğru pozitif oranını göstermektedir. Grafikte, tüm sınıfların AUC değerleri belirtilerek "Epik Şiir" için AUC değeri 0.90 ile en yüksek performans görülmüştür.



Burada, modelin test veri seti üzerindeki tahmin performansını karmaşıklık matrisi aracılığıyla göstermekteyiz. Matriste, doğru tahmin edilen sınıflar ve yanlış tahmin edilen sınıflar yer almaktadır. Örneğin, "Epik şiir" için model 823 doğru tahmin yaparken 44 yanlış tahmin gerçekleştirmiştir. Bu görsel, modelin her sınıf için güçlü ve zayıf yönlerini detaylı şekilde analiz etmek için kullanılmıştır.

**Accuracy:** Modelin genel doğruluğu 0.79'dur. Bu, modelin tüm veri setindeki sınıfları doğru bir şekilde tahmin etme oranını ifade etmektedir.

**Recall :** Satirik Şiir için en yüksek recall değeri 0.81 olarak gözlemlenmiştir. En düşük recall değeri ise Pastoral Şiir için 0.78'dir.

**Precision:** En yüksek precision değeri Epik Şiir ve Pastoral Şiir için 0.87 olarak görülmüştür. En düşük precision değeri ise Lirik Şiir için 0.68 olarak kaydedilmiştir.

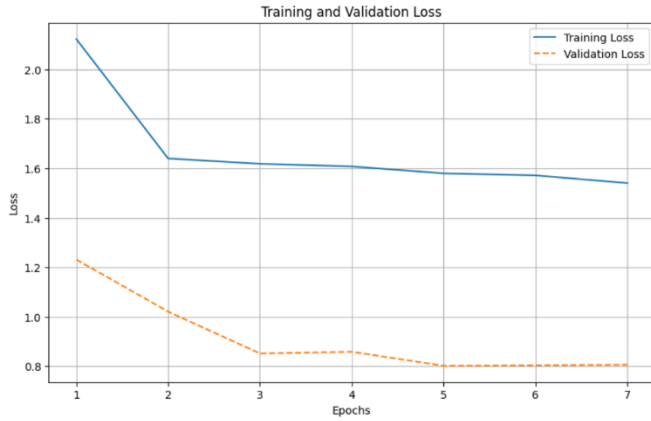
**Sensitivity:** Satirik Şiir için en yüksek duyarlılık değeri 0.81 ile başarılı bir sonuç elde edilmiştir.

**Specifity:** Epik Şiir ve Pastoral Şiir için özgüllük değerleri 0.97 olarak yüksek çıktı.

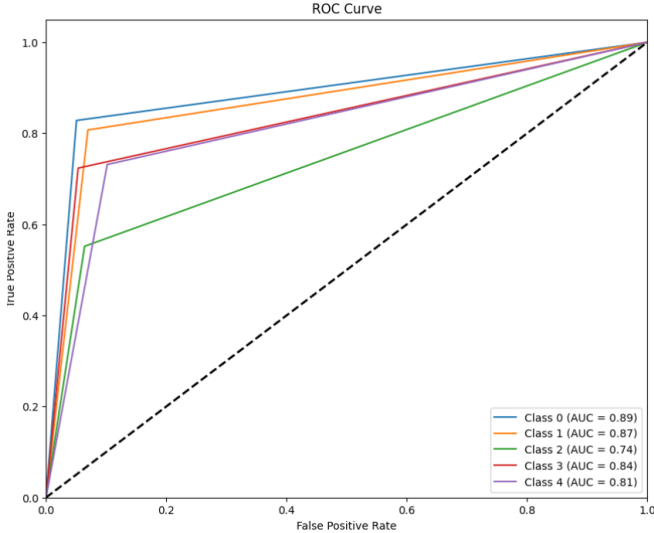
**F-Score:** Epik Şiir ve Satirik Şiir en yüksek F1-score değerlerine sahip olup sırasıyla 0.85 ve 0.80 olarak görülmüştür.

## B. GPT-2

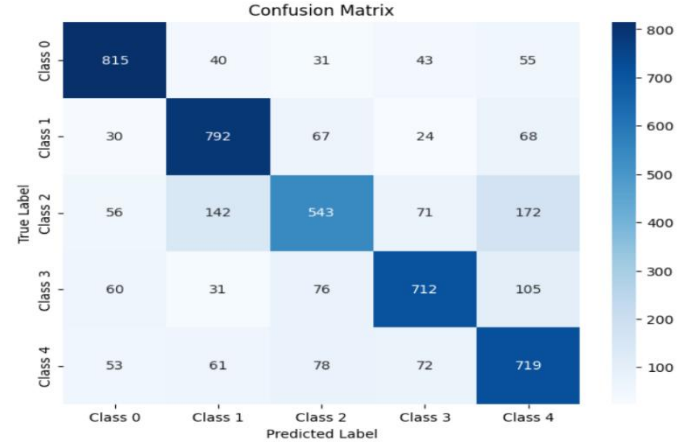
GPT-2 modeli, sınıflandırma görevinde güçlü bir dil işleme kapasitesi ile ön plana çıkmaktadır. Bu model de eğitim sırasında Early Stopping gibi optimizasyon teknikleriyle desteklenmiş, her epoch'ta modelin doğruluk ve kayıp değerleri yakından izlenmiştir. Metinlerin maksimum 512 token uzunluğunda işlenmesi sağlanmıştır. Eğitim sonunda model, 7 epoch'ta optimum performansa ulaşmıştır. En düşük doğrulama kaybına sahip model kaydedilmiştir.



Grafikte, eğitim kaybının her epoch sonunda azaldığı ve modelin öğrenme sürecinin devam ettiği gözlemlenmiştir. Doğrulama kaybı da benzer şekilde azalarak modelin genelleştirme yeteneğinin arttığı görülmektedir.



Bu grafik, modelin sınıf bazında ROC eğrilerini ve her sınıf için AUC değerlerini göstermektedir. ROC eğrisi, modelin doğru pozitif oranına karşılık yanlış pozitif oranını göstermektedir. Bu değerler, modelin özellikle Epik Şiir ve Pastoral Şiir sınıflarında daha yüksek doğruluk sağladığını göstermektedir.



Bu grafik, Epik Şiir için model 815 doğru tahmin yaparken, 43 örneği Satirik Şiir olarak yanlış sınıflandırmıştır. Grafik, modelin güçlü yönlerini ve iyileştirilmesi gereken sınıfları detaylı bir şekilde analiz etmeyi sağlamaktadır.

**Accuracy:** Modelin genel doğruluğu %73 olarak hesaplanmıştır.

**Recall :** Epik Şiir ile en yüksek recall değeri 0.83 olarak elde edilmiştir.

En düşük recall değeri ise Lirik Şiir için 0.55 olarak kaydedilmiştir.

**Precision:** En yüksek precision değeri Epik Şiir için 0.80 olarak gözlemlenmiştir. En düşük precision değeri ise Felsefi Şiir için 0.64 olarak kaydedilmiştir.

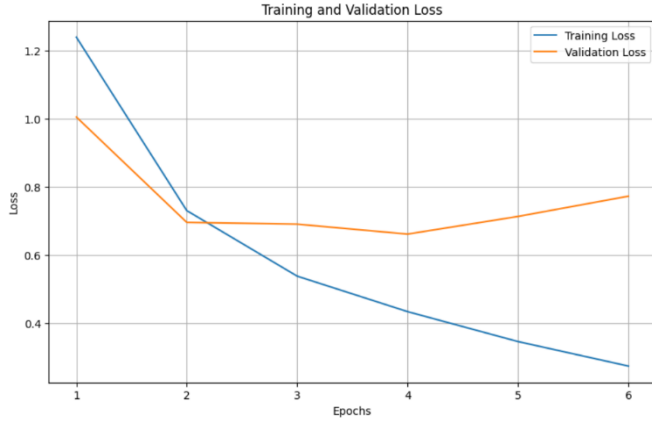
**Sensitivity:** Epik Şiir de 0.83 değeri görülerek bu sınıfta model, pozitif örneklerin %83'ünü doğru bir şekilde tespit edebilmiştir.

**Specificity:** Epik Şiir de 0.95 ile model, bu sınıfa ait olmayan örnekleri %95 oranında doğru bir şekilde ayırt edebilmiştir.

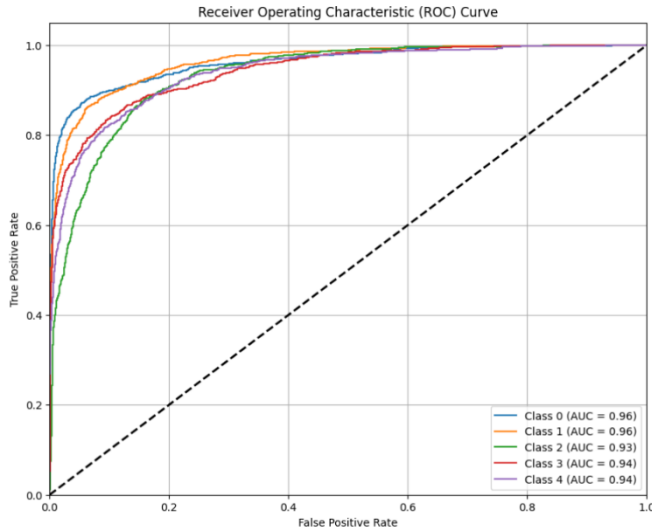
**F-Score:** En yüksek F1-score değerleri Epik Şiir için 0.82 ve Pastoral Şiir için 0.77 olarak hesaplanmıştır.

### C. ELECTRA

Electra, verimli ve hızlı bir şekilde token düzeyinde sınıflandırma yapmak için geliştirilmiş bir transformer modelidir. Bu çalışmada "electra\_turkish\_poetry\_classifier" isimli model kullanılmıştır. Electra'nın tokenizer'ı kullanılarak metinler input\_ids ve attention\_mask formatında dönüştürülmüştür. Eğitim ve çıkarım işlemleri GPU destekli olarak gerçekleştirilmiştir.



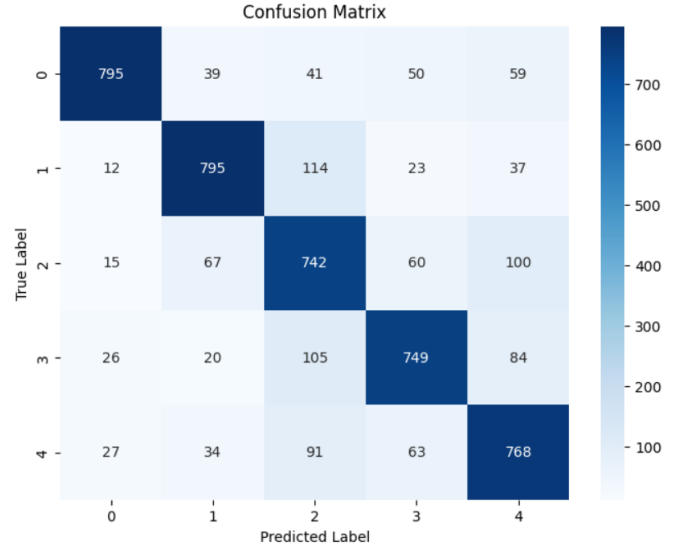
Eğitim kaybının sürekli azaldığı, doğrulama kaybının ise bir noktadan sonra az da olsa artmaya başladığı gözlemlenmiştir.



Grafikte her sınıfa ait ROC eğrileri farklı renklerle gösterilmiştir. Tüm sınıflar için AUC değerleri 0.93 ile 0.96 arasında değişmektedir, bu da modelin güçlü bir performansa sahip olduğunu bize göstermektedir.

En yüksek performans Epik Şiir ve Pastoral Şiir'de gözlemlenmiştir.

Lirik Şiir, diğer sınıflara göre nispeten daha düşük bir AUC değerine sahiptir. Eğrilerin sol üst köşeye yakınlığı, modelin yüksek duyarlılığa sahip olduğunu gösterir.



Satırlar gerçek etiketleri, sütunlar modelin tahmin ettiği etiketleri temsil etmektedir. Diyagonal boyunca yer alan değerler, doğru sınıflandırılmış örneklerin sayısını ifade eder. Epik Şiir ve Pastoral Şiir için oldukça başarılı sonuçlar alınmıştır. Bu sınıflar için doğru sınıflandırma oranı yüksektir. Diyagonal değerlerin yüksek olması, modelin genel doğruluğunun iyi olduğunu göstermektedir. Modelin test verisindeki performans ölçümleri genel olarak bu şekildedir:

**Accuracy : %78.29**

**Precision : %78.90**

**Recall : %78.29**

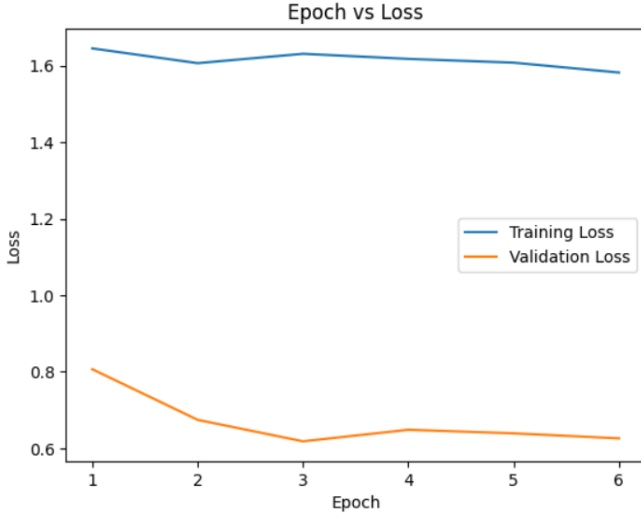
**F-Score: %78.47**

**Specificity : %95.32**

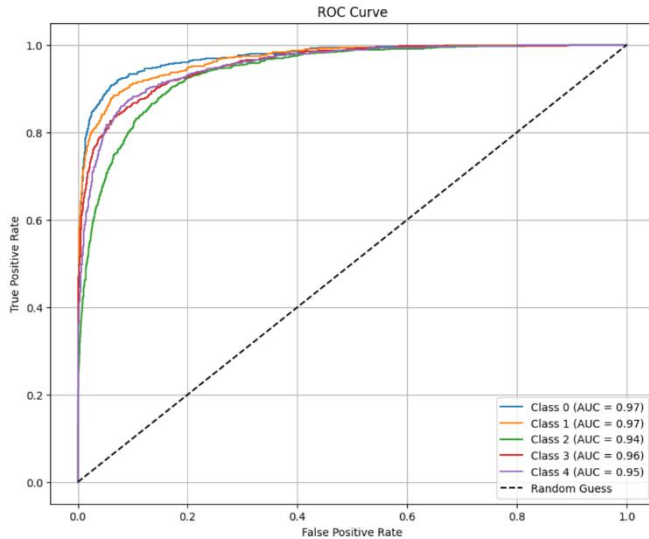
**AUC: %86.43**

#### D. ROBERTA

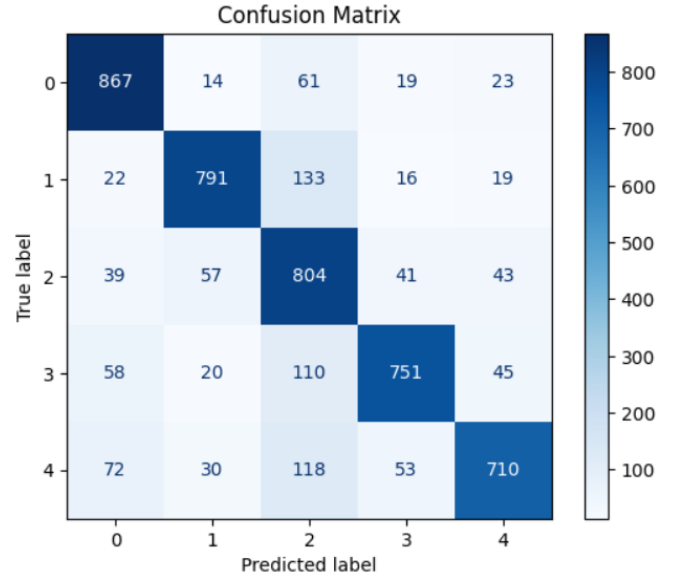
Kullanılan modelin adı "xlm-roberta-base" dir. Model Tipi: XLM-Roberta tabanlı çok dilli dil modeli. Veri setlerinde her bir satır bir şiiri ve karşılık gelen etiket bilgisini içerir. Metinlerin ortalama uzunluğu analiz edildi ve tokenizasyon sırasında bilgi kaybını minimize etmek için dikkat edildi. Sınıfların dengeli olmasına özen gösterildi.



Eğitim kaybı başlangıçta yüksek bir değerle başlayıp her epoch ilerledikçe azalma göstermiştir. Doğrulama kaybı başlangıçta hızlı bir şekilde düşüş göstermiş ve daha sonra stabil bir seviyeye ulaşmıştır.



Tüm sınıflar için ROC eğrileri, rastgele tahmin çizgisinden oldukça uzakta yer almaktadır. Epik Şiir ve Pastoral Şiir en yüksek AUC değerine sahiptir (0.97). Bu sınıflar için modelin tahmin performansı oldukça iyidir.



Epik Şiir ve Satirik Şiir gibi sınıflar için yüksek doğruluk oranları elde edilmiştir.

**Accuracy:** Modelin, tüm sınıflar genelinde doğru sınıflandırma oranı %80'dir.

**Recall :** Model, Epik Şiir 'ı doğru bir şekilde tespit etmede en başarılı olmuştur.

**Precision:** Modelin, Pastoral Şiir için yaptığı tahminlerde en doğru sonuçları elde ettiği gözlemlenmiştir.

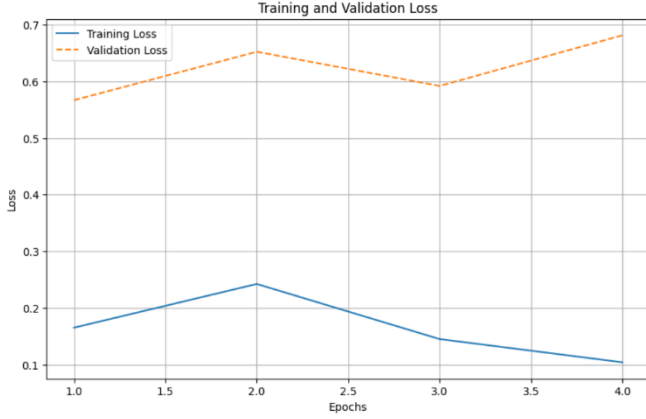
**Sensitivity:** Model, Epik Şiir için yüksek duyarlılık göstermiştir. Felsefi Şiir için ise duyarlılık, diğer sınıflara göre düşük olarak gözlemlenmiştir.

**Specificity:** Model, specificity açısından tüm sınıflarda oldukça iyi performans göstermektedir.

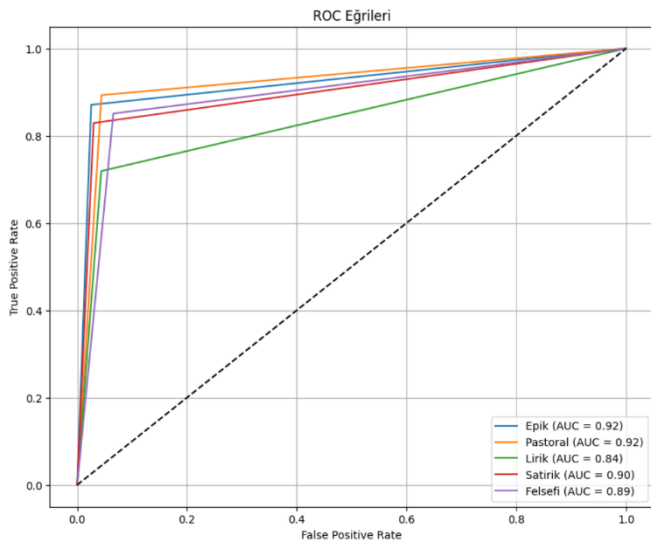
**F-Score:** Precision ve Recall dengesine göre en iyi performans Epik Şiir 'dedir.

### E. DistilBERT

Model adı: distilbert-base-multilingual-cased dir. Her satırda bir şiir metni ve karşılık gelen sınıf etiketi (label) bulunmaktadır. Veri setindeki tüm sınıflar dengeli bir şekilde dağıtılmıştır. DistilBERT tokenizer, şiir metinlerini 512 token'a kadar kırpma ve doldurma işlemleriyle hazırlar.

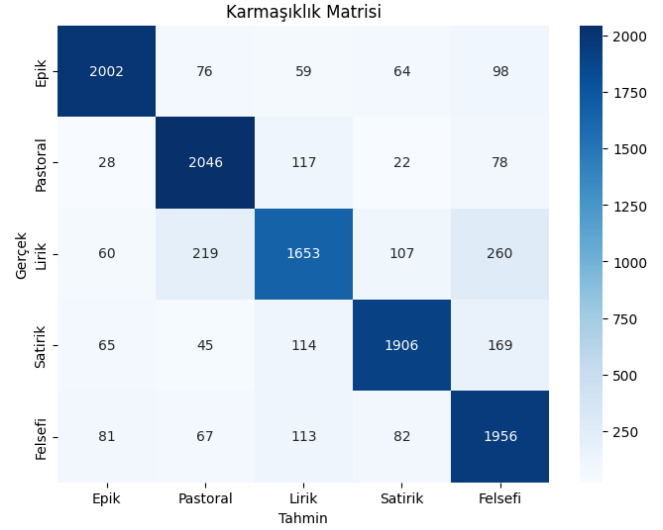


Eğitim kayıplarının sürekli olarak azaldığını görmekteyiz, bu da modelin eğitim verisi üzerinde öğrenmeye devam ettiğini göstermektedir. Epoch 2'den sonra önemli bir düşüş görülmüş ve son epoch'ta minimum değere ulaşılmıştır. Eğitim ve doğrulama kayıplarının birbirine yakın olması, modelin iyi bir genel performansa sahip olduğunu işaret etmektedir.



Epik ve Pastoral sınıflar için AUC değeri 0.92 ile en yüksek performansı göstermektedir. Model, bu sınıflar için yüksek ayırım gücüne sahiptir. Lirik Şiir sınıfı AUC değeri 0.84 ile en düşük performansı sergilemektedir. Bu durum, modelin

bu sınıfı diğerlerinden ayırt etmede zorlandığını gösterir.



En iyi performans Epik ve Pastoral sınıflarında gözlemlenmiştir. Bu sınıflar için doğru sınıflandırma oranı oldukça yüksek çıkmıştır. Lirik sınıfında 260 örnek Felsefi olarak yanlış sınıflandırılmıştır. Bu durum, modelin bu iki sınıf arasındaki ayrımı yapmakta zorlandığını göstermektedir.

**Accuracy:** Modelin genel doğruluk oranı 0.8325 şeklinde çıkmıştır, test veri setindeki örneklerin %83.25'inin doğru sınıflandırıldığını göstermektedir.

**Recall :** Gerçek pozitif sınıfların %83.25'i model tarafından doğru olarak tespit edilmiştir.

**Precision:** Modelin tahmin ettiği pozitif sınıflardan %83.42'si doğru olarak tahmin edilmiştir.

**Sensitivity:** Epik Şiir, sensitivity %87.08 açısından en iyi performansı göstermiştir. Lirik Şiir, sensitivity açısından en düşük değeri %71.90 almıştır.

**Specifity:** Tüm sınıflar için specificity değerlerinin %93 üzerinde olması, modelin genel olarak sınıfları birbirinden doğru bir şekilde ayırt ettiğini göstermektedir.

**F-Score:** 0.83 şeklinde sonuçlanmıştır. Precision ve Recall değerlerinin harmonik ortalaması, modelin genel performansını dengeli bir şekilde göstermektedir.

### III. MODELLERİN KARŞILAŞTIRILMASI

#### 1. DistilBERT:

Genel doğruluk oranı %83.25 ile en iyi performansı sergilemiştir. ROC eğrisi, tüm sınıflarda dengeli sonuçlar sunmuş ve özellikle Epik ve Pastoral Şiir sınıflarında yüksek AUC değerleri elde edilmiştir.

#### 2. ROBERTA:

ROBERTA, %80 doğruluk oranıyla DistilBERT'in hemen arkasından gelmektedir. Epik ve Pastoral Şiir sınıflarında 0.97 AUC ile sınıf bazında en iyi ayrım gücünü göstermiştir. Eğitim ve doğrulama kayıpları dengeli bir şekilde azalmıştır.

#### 3. BERT:

BERT %79 doğruluk oranıyla genel performansta iyi bir sonuç sunmuştur. Ancak, sınıf bazında performans ROBERTA ve DistilBERT'in biraz gerisinde kalmıştır.

#### 4. ELECTRA:

ELECTRA, %78.29 doğruluk oranı ve genel AUC (%86.43) ile güçlü bir model olmasına rağmen, diğer modellere göre dengeli performansta biraz geride kalmıştır. Özellikle Lirik Şiir sınıfında düşük sonuçlar gözlenmiştir.

#### 5. GPT-2:

GPT-2 %73 doğruluk oranıyla en düşük performansı göstermiştir. Özellikle Lirik Şiir sınıfında %55 sensitivity ile başarısız olmuştur. Eğitim sürecinde öğrenme daha yavaş gerçekleşmiştir.

### IV. REFERENCES

- [1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need. In Advances in Neural Information Processing Systems (NIPS).
- [2] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.
- [3] Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., & Brew, J. (2020). Transformers: State-of-the-Art Natural Language Processing. EMNLP.
- [4] Chatgpt.com
- [5] Google Colab: Research Platform for Machine Learning.
- [6] <https://antoloji.com/>
- [7] Python Software Foundation. Python Language Reference.

