### **3.1.1.3 Long Short-Term Memory (LSTM) ile Müzik Türü Sınıflandırma**

**LSTM (Long Short-Term Memory):** Ses verisindeki zamana bağlı değişimleri ve sıralı kalıpları yakalamak üzere tasarlanmıştır. Özellikle ritim ve melodi gibi zamanla evrilen müzikal öğelerin analizinde öne çıkar.

**1. LSTM Nedir?**

Uzun Kısa Süreli Bellek (LSTM), Tekrarlayan Sinir Ağlarının (RNN) gelişmiş bir varyasyonudur. Geleneksel RNN yapılarının uzun sekanslarda karşılaştığı gradyan kaybolması (vanishing gradient) sorununa çözüm getirmek amacıyla geliştirilmiştir. LSTM ağları, sahip oldukları hücre durumu (cell state) ve özel kapı mekanizmaları (unutma, girdi ve çıktı kapıları) aracılığıyla bilgiyi uzun süre hafızada tutabilir ve hangi bilginin korunup hangisinin atılacağını veya çıktıya dönüştürüleceğini etkin bir şekilde yönetebilir. Bu yetenekleri sayesinde, sıralı verilerdeki karmaşık zamansal ilişkileri modellemede oldukça başarılıdırlar.

**2. Müzik Türü Sınıflandırmada LSTM Kullanımı**

Müziğin yapısı temel olarak zamansaldır; melodi, armoni ve ritim gibi öğeler zaman ekseninde bir sıra takip eder. LSTM modelleri, müzik sinyallerinden elde edilen zaman serisi formatındaki özellikleri (örneğin, zaman içindeki MFCC veya kromagram değerleri) işleyerek bu zamansal örüntüleri öğrenebilir. Belirli bir zaman dilimindeki ses özelliklerinin, kendinden önceki ve sonraki özelliklerle olan ilişkisini modelleyerek müzik türünü tahmin etmeye çalışır. Uygulamada genellikle ses dosyası kısa süreli pencerelere ayrılır ve bu pencerelerden çıkarılan özellik dizileri LSTM modeline girdi olarak sunulur.

**3. Literatürde LSTM Uygulamaları**

Müzik türü sınıflandırması üzerine yapılan akademik çalışmalarda LSTM kullanımına dair bazı örnekler aşağıda sunulmuştur:

* **Sudarsanam, vd. (2021):** "Bu makale, müzik türlerini sınıflandırmak için LSTM ve Evrişimli Tekrarlayan Sinir Ağları (CRNN) başta olmak üzere derin öğrenme modellerini karşılaştırmaktadır. Bulgular, LSTM temelli yaklaşımların müzikteki zamansal nitelikleri kavramada ve tür sınıflandırmasında yüksek başarım göstermede etkili olduğunu ortaya koymaktadır." [1]
* **Liu, vd. (2020):** "Çalışmada, MFCC özelliklerini girdi olarak alan bir LSTM ağı ile müzik türü sınıflandırması gerçekleştirilmiştir. Geliştirilen model, farklı müzik türlerine özgü zamansal kalıpları öğrenerek GTZAN veri kümesinde dikkate değer sonuçlar vermiştir." [2]
* **Singh, vd. (2019):** "Bu araştırmada, müzik türü sınıflandırması amacıyla CNN ve LSTM modellerini birleştiren hibrit bir mimari önerilmiştir. CNN spektrogramlardan yerel öznitelikleri çıkarırken, LSTM bu özniteliklerin zamansal akışını modellemiştir. Önerilen hibrit modelin, tekil modellere kıyasla daha iyi performans sergilediği gözlemlenmiştir." [3]

**4. LSTM Modelinin Yapısı**

Müzik türü sınıflandırması görevine yönelik tipik bir LSTM modeli genellikle şu katmanları içerir:

* **Girdi Katmanı:** Sıralı ses öznitelikleri (örn. MFCC dizileri gibi zaman serileri).
* **LSTM Katmanları:** Bir veya daha fazla sayıda LSTM katmanı, girdi dizisindeki zamansal bağımlılıkları işler.
* **Tam Bağlantılı Katmanlar (Fully Connected Layers):** LSTM katmanlarından elde edilen temsili sınıflandırma için uygun formata getirir.
* **Çıkış Katmanı:** Genellikle Softmax aktivasyon fonksiyonu kullanılarak her bir müzik türü için bir olasılık değeri hesaplar.

**5. Avantajları ve Sınırlamaları**

**Güçlü Yönleri:**

* Veri içindeki uzun süreli zamansal ilişkileri öğrenme kapasitesi (Bu, özellikle MLP gibi sıralı bağımlılıkları doğrudan modellemeyen veya XGBoost gibi genellikle zamansal olmayan özellik setleri üzerinde çalışan modellere göre belirgin bir avantajdır).
* Sıralı verilerdeki (örn. zaman serisi halindeki ses özellikleri) desenleri tespit etmede yüksek etkinlik.
* Gradyan kaybolması problemine karşı standart RNN'lere göre daha dirençli olması.

**Zayıf Yönleri:**

* Tekrarlayan yapısı nedeniyle, genellikle daha yüksek hesaplama maliyeti gerektirebilir (Özellikle daha basit MLP yapılarına veya bazı optimize edilmiş ağaç tabanlı modellere kıyasla).
* Özellikle çok uzun dizilerin eğitimi daha fazla zaman gerektirebilir.
* Model karmaşıklığından dolayı en uygun hiperparametrelerin bulunması deneme yanılma gerektirebilir.

**6. Sonuç**

LSTM ağları, müziğinzamansal karakteristiğinianlama ve modellemedeki yetenekleri ile müzik türü sınıflandırma probleminde değerli bir yaklaşım sunar. Projede kullanılan diğer modellerden (CNN'in görsel/uzamsal örüntüleri yakalaması, MLP veya XGBoost'un farklı özellik temsilleri üzerinden sınıflandırma yapması gibi) farklı olarak, LSTM'in temel gücü sıralı bağımlılıkları ve zaman içindeki dinamikleri kavramaktır. Özellikle ses sinyallerindeki sıralı yapının ve uzun erimli bağımlılıkların kritik olduğu senaryolarda başarılı sonuçlar üretebilirler. Ayrıca, CNN gibi farklı mimarilerle entegre edilerek hibrit modeller oluşturulması da mümkündür.

**Kaynaklar:**

[1] Sudarsanam, S., et al. (2021). Deep Learning Models for Music Genre Classification: A Comparative Study. International Journal of Advanced Computer Science and Applications (IJACSA), 12(5).

[2] Liu, Y., et al. (2020). Music Genre Classification Based on LSTM Network Using MFCC Features. Applied Sciences, 10(14), 4973.

[3] Singh, S., et al. (2019). Music Genre Classification using Hybrid CNN-LSTM Model. 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS).