Data Analytics for Artificial Intelligence: Göreceğimiz Konular  
[*http://courses.ncirl.ie/index.cfm/page/module/moduleId/68622*](http://courses.ncirl.ie/index.cfm/page/module/moduleId/68622)

AI geliştiricisi veya Makine Öğrenimi (ML) geliştiricisi olmak için aşağıdaki konular özellikle önemlidir:

1. **Data Collection and Data Manipulation**:
   * **Veri Kaynakları**: Farklı veri kaynaklarından veri toplama ve entegrasyon.
   * **Veri İşleme**: Sütun ve satır seçimi, gruplama, birleştirme, filtreleme.
2. **Data Preprocessing and Transformation Strategies**:
   * **Veri Temizleme**: Eksik değerlerin ele alınması, aykırı değerlerin filtrelenmesi.
   * **Özellik Ölçeklendirme**: Normalizasyon, standardizasyon.
   * **Özellik Seçimi**: Boyut azaltma yöntemleri ve önemli özelliklerin belirlenmesi.
3. **Statistical Analysis - Hypothesis & Inference**:
   * **Hipotez Testleri**: T-Testi, ANOVA, ki-kare testleri.
   * **Korelasyon ve Regresyon Analizi**: Değişkenler arasındaki ilişkilerin incelenmesi.
4. **Dimension Reduction Methods**:
   * **Principal Component Analysis (PCA)**: Veri setinin boyutunu azaltırken bilgi kaybını minimize etme.
   * **Faktör Analizi**: Gizli değişkenlerin keşfi.
5. **Prediction (Regression)**:
   * **Basit ve Çoklu Doğrusal Regresyon**: Sürekli değişkenlerin tahmini.
   * **Lojistik Regresyon**: İkili sınıflandırma problemleri için kullanılır.
6. **Classification**:
   * **K-En Yakın Komşu (K-NN)**: Basit ve etkili bir sınıflandırma algoritması.
   * **Karar Ağaçları ve Rastgele Ormanlar**: Karar verme süreçlerini modelleme.
   * **Destek Vektör Makineleri (SVM)**: Yüksek boyutlu verilerde etkili sınıflandırma.
7. **Clustering**:
   * **K-Means Clustering**: Veri noktalarını benzerliklerine göre gruplandırma.
   * **Hierarchical Clustering**: Veri kümeleri arasındaki hiyerarşik ilişkileri belirleme.
   * **Density-Based Clustering (DBScan)**: Karmaşık veri yapılarında kümeleri keşfetme.
8. **Modelling and Evaluation**:
   * **Model Doğrulama**: Eğitim, test ve doğrulama veri setlerinin kullanımı.
   * **Performans Metrikleri**: Doğruluk, kesinlik, geri çağırma, F1 skoru.
   * **Çapraz Doğrulama**: Modelin genelleme yeteneğinin değerlendirilmesi.
9. **Content Analysis**:
   * **Doğal Dil İşleme (NLP)**: Metin verilerinin analizi ve işlenmesi.
   * **Metin Madenciliği**: Önemli bilgilerin ve kalıpların çıkarılması.
   * **Kelime Gömme Vektörleri**: Kelimelerin sayısal temsil edilmesi (örneğin, Word2Vec, GloVe).
10. **Programming Skills and Tools**:
    * **Python veya R**: Veri bilimi ve ML için yaygın olarak kullanılan programlama dilleri.
    * **Makine Öğrenimi Kütüphaneleri**: TensorFlow, PyTorch, scikit-learn gibi araçların kullanımı.

**Sonuç olarak**, AI ve ML geliştiricisi olmak için güçlü bir istatistik ve matematik temeline sahip olmanız, veri işleme ve modelleme konularında yetkin olmanız gerekmektedir. Ayrıca, algoritmaların teorik temellerini anlamak ve bunları pratik uygulamalarla pekiştirmek önemlidir.

Data Analytics for Artificial Intelligence (English Notes):