

Sözcüksel Analiz ve Tokenleştirme Nedir?

Çalışma Kağıdı

Sözcüksel analiz, bir metni veya program kodunu anlamlı birimlere (token'lara) ayırma işlemidir. Tokenleştirme ise bu birimlere ayırma eyleminin kendisidir ve genellikle boşluklar, noktalama işaretleri veya belirli kurallar kullanılarak yapılır.

Sorular

1. Aşağıdakilerden hangisi tokenleştirme için bir ayırıcı *olamaz*?

- A) Boşluk
- B) Noktalama işareti
- C) Harf
- D) Özel karakter

2. Bir derleyicinin ilk aşaması genellikle nedir?

- A) Sözcüksel Analiz
- B) Yazım Denetimi
- C) Görsel Tasarım
- D) Veritabanı Yönetimi

3. "Bilgisayar mühendisliği" ifadesi, boşluk ayırıcısı ile tokenleştirildiğinde kaç token oluşur?

- A) 1
- B) 2
- C) 3
- D) 4

4. Basit bir cümle için tokenleştirme örneği:

5. Bir program kod satırı için tokenleştirme örneği:

6. Farklı ayırıcılar kullanarak tokenleştirme:

7. Tanımla: Sözcüksel Analiz Nedir?

8. Tanımla: Token Nedir?

9. Tanımla: Tokenleştirme Nasıl Yapılır?

Cevap Anahtarı

1. C) Harf — Harfler genellikle token'ların kendisini oluşturur, ayırıcı olarak kullanılmazlar. Boşluk, noktalama ve özel karakterler ise token'ları ayırmak için kullanılır.
2. A) Sözcüksel Analiz — Derleyicilerde ilk aşama, kaynak kodu token'lara ayırmak olan sözcüksel analizdir.
3. B) 2 — Boşluk karakteri, 'Bilgisayar' ve 'mühendisliği' kelimelerini ayırır, bu da 2 token anlamına gelir.
4. Cümle: "Merhaba dünya! Bugün hava güzel." Token'lar: ["Merhaba", "dünya", "!", "Bugün", "hava", "güzel", "."]
5. Kod satırı: `int sayi = 10;` Token'lar: ["int", "sayi", "=", "10", ";"]
6. Metin: "elma-armut_kiraz" Ayırıcılar: '-', '_' Token'lar: ["elma", "armut", "kiraz"]
7. Bir metni veya kodu anlamlı birimlere (token'lara) ayırma süreci.
8. Sözcüksel analiz sonucunda elde edilen anlamlı en küçük birim.
9. Genellikle boşluklar, noktalama işaretleri veya önceden tanımlanmış kurallar kullanılarak.

Bounlu

Tüm kartlar, adım adım çözümler ve AI hoca desteği Notek uygulamasında.
Sınav tarihlerini Promy otomatik hatırlatıcıya çevirir.