

METİN SINIFLANDIRMA KILAVUZU

1. Giriş

Metin sınıflandırma, doğal dil işlemedeki pek çok problemin temelini oluşturuyor ve pek çok uygulamada hayat buluyor. Örneğin mesajları gerekli / gereksiz-spam olarak sınıflandıran programlar; verilen bir metnin olumlu mu olumsuz mu veya üzgün mü, neşeli mi, nötr mü olduğunu bulan duygu analizi yöntemleri; öneri sistemleri; dil tanıma araçları, en yaygın kullanım alanları. Metinlerin akıcılığı, zorluğu, okunabilirliğini ölçen veya bir metni kimin yazdığını, yazar kişinin metnin dilini anadil mi, yabancı dil olarak mı bildiğini bulan, metinlerdeki argümanları çıkarıp taraflardan hangisinin haklı olduğunu tespit etmeye çalışan algoritmalar da son yıllarda üzerinde çalışılan ilginç sınıflandırma problemlerinden¹. Bu yazıda sizlere cognitius metin sınıflandırma özelliklerini kullanarak nasıl çalışılabileceğini gösterelim.

2. Sisteme Giriş ve Modül Oluşturma

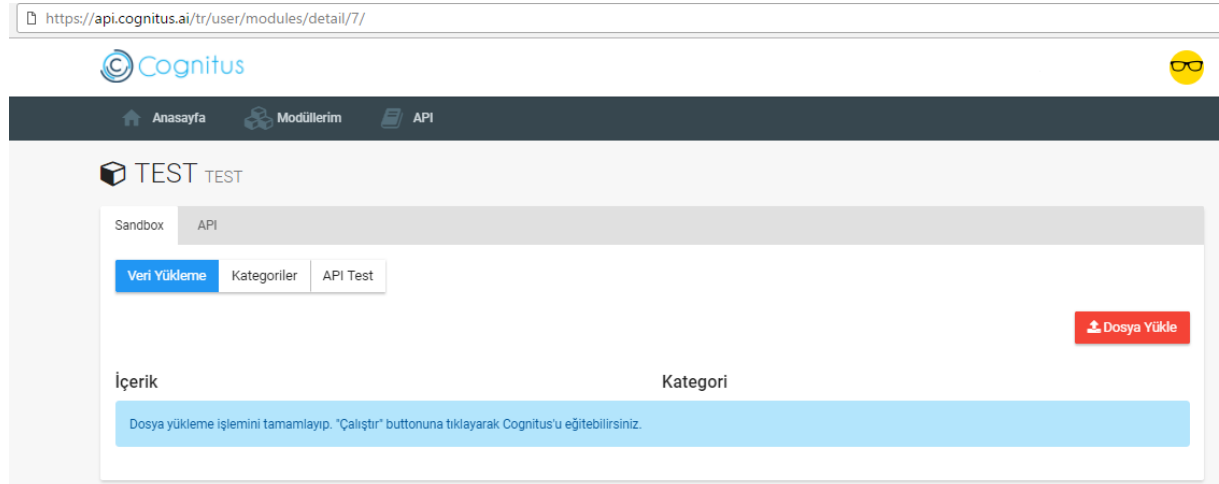
Sisteme giriş <http://api.cognitus.ai/> linkinden yapıldıktan sonra kullanıcı hesabı oluşturulur. Kayıt işlemi tamamlandıktan sonra sisteme giriş yapılır. Giriş sayfasında kullanıcı tarafından oluşturulan modüller listelenir (ör: “TEST” modülü). “Modüllerim” bölümündeki “Yeni Modül” sekmesinden kullanıcı tarafından modül oluşturulabilmektedir:



Şimdi yeni bir modül oluşturalım. “İsim” sekmesine modül ismi, “Açıklama” sekmesine modül açıklaması yazılır ve “Dil” sekmesinden ise dil seçeneği ayarlanır:

¹ <http://cognitus.ai/yazi/metin-siniflandirma-ne-ise-yarar/>

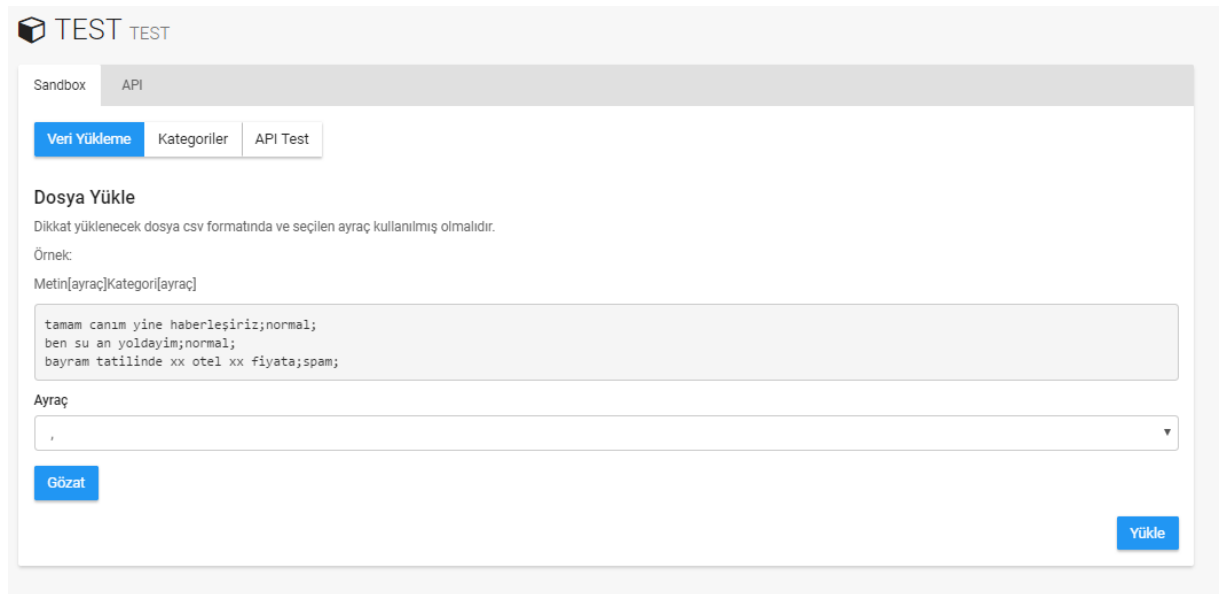
Yeni modül oluşturulduktan sonra aşağıdaki gibi boş modül elde edilir:



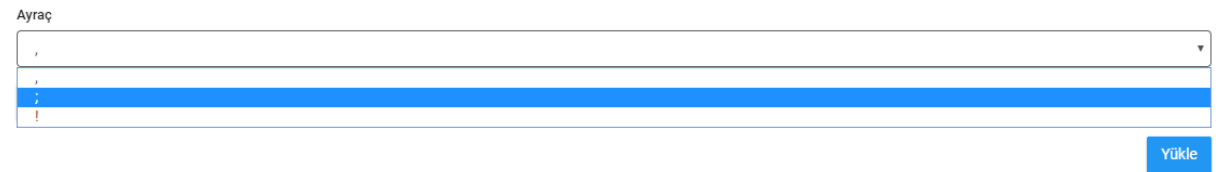
Buradaki url' deki (<https://api.cognitus.ai/tr/user/modules/detail/7/>) 7 sayısı "module_id" yi referans etmektedir.

3. Modüle Veri Yükleme

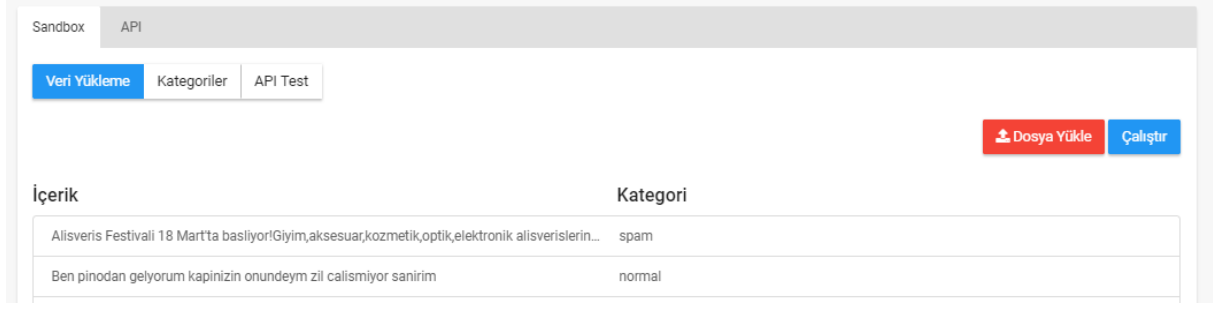
Bu bölümde "Dosya Yükle" butonu ile sisteme sınıflandırılacak dosya yüklenir. Açılan pencereden "Gözet" butonu ile yüklenecek dosya seçilir:



"Ayraç" bölümünden kategori ve içerik ayrımının nasıl yapılacağı seçilir ve "Yükle" butonu ile veriler sisteme yüklenir:



Örnek olarak [buradan](#) alınan veriler kullanılmıştır. Veriler yüklendikten sonra aşağıdaki gibi listelenmektedir ve “Çalıştır” butonu ile sistem eğitilir:

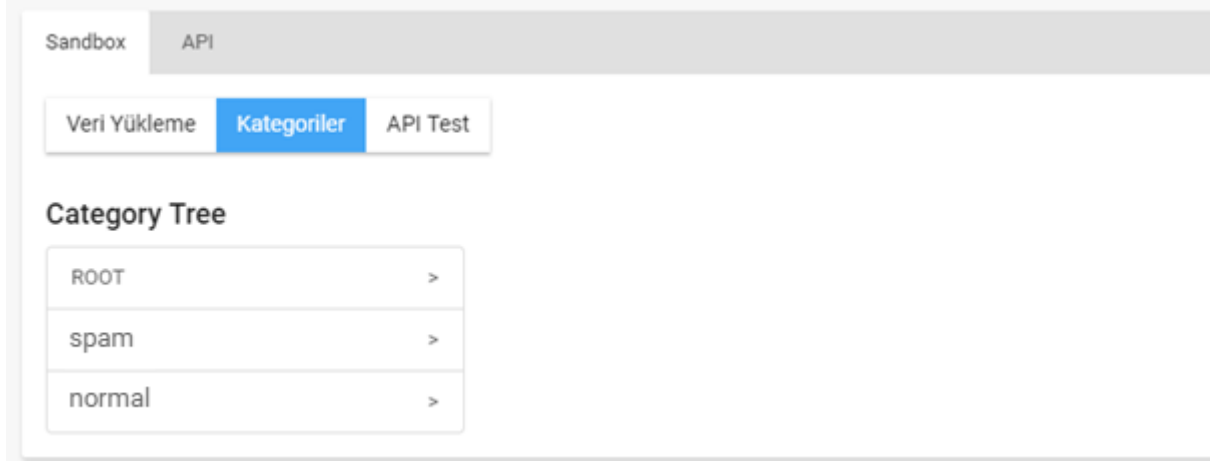


The screenshot shows a web interface with a header bar containing 'Sandbox' and 'API'. Below the header, there are three tabs: 'Veri Yükleme', 'Kategoriler', and 'API Test'. On the right side, there are two buttons: 'Dosya Yükle' (red) and 'Çalıştır' (blue). The main content area displays a table with two columns: 'İçerik' and 'Kategori'. The table contains two rows of data.

İçerik	Kategori
Alisveris Festivali 18 Mart'ta basılıyor!Giyim,aksesuar,kozmetik,optik,elektronik alisverislerin...	spam
Ben pinodan gelyorum kapinizin onundeym zil calismiyor sanirim	normal

4. Veri Analizi

Eğitim tamamlandıktan sonra veri analizi için “Kategoriler” sekmesine tıklanır ve “Category Tree” gözlemlenir:



The screenshot shows the same web interface as before, but the 'Kategoriler' tab is selected. Below the tabs, there is a section titled 'Category Tree' which displays a list of categories: 'ROOT', 'spam', and 'normal'. Each category has a right-pointing arrow next to it.

Category Tree
ROOT >
spam >
normal >

Kullandığımız örnekteki veriler “spam” ve “normal” olmak üzere 2 kategoriye ayrılmıştır. Bundan dolayı “Category Tree” de ROOT altında iki kategori oluşmuştur.

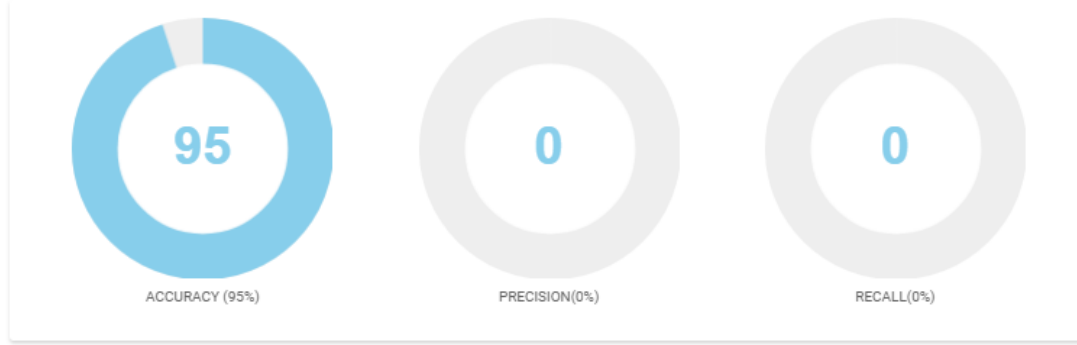
4.1. ROOT Kategorisi

Sistem tarafından her zaman oluşturulan “ROOT” kategorisi eğitim hakkında genel bilgi verir:

Samples



Statics



Keywords

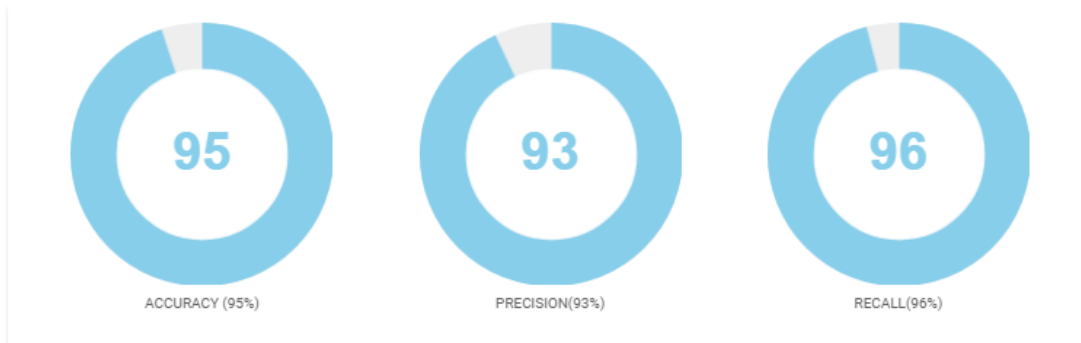
4.2. Diğer Kategoriler

Her eğitimden sonraki kategoriler “ROOT” altında listelenir ve burada kategori ile ilgili sample, statics, keyword bilgileri gösterilir. Örneğin “spam” kategorisinin bilgileri aşağıdaki gibidir:

Samples



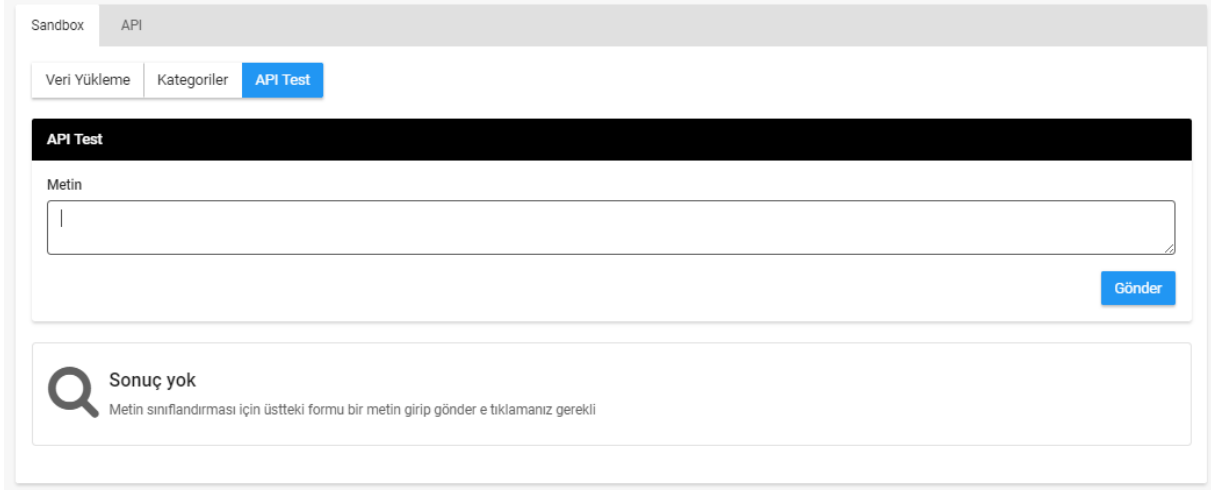
Statics



Keywords

5. API Test

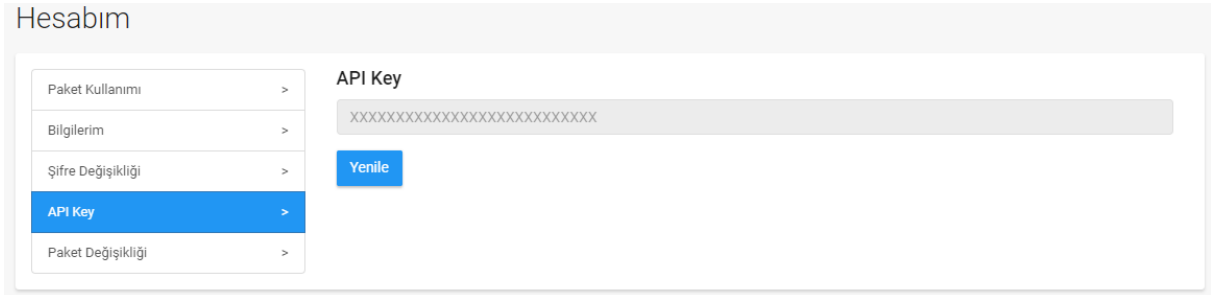
Eđitim iřlemi tamamlandıktan sonra test ařamasına geçilir. Burada sınıflandırılacak metin, “Metin” sekmesine yazılır ve sonuçlar gözlemlenir:



The screenshot shows a web interface for API testing. At the top, there are tabs for "Sandbox" and "API". Below the tabs, there are three buttons: "Veri Yükleme", "Kategoriler", and "API Test". The "API Test" button is highlighted. Below the buttons, there is a section titled "API Test" with a sub-section "Metin". There is a text input field with a cursor and a "Gönder" button. Below the input field, there is a search icon and the text "Sonuç yok" (No results). Below that, there is a message: "Metin sınıflandırması için üstteki formu bir metin girip gönder e tıklamanız gerekli" (Text classification requires you to enter a text and click send in the form above).

6. İstek Göndererek Kullanım

Kullanıcı sisteme giriş yaptıktan sonra “Hesabım” bölümüne gelip kendisi için tanımlanan token’i “API Key” içerisinden görebilir:



The screenshot shows the "Hesabım" (My Account) page. On the left, there is a sidebar menu with the following items: "Paket Kullanımı", "Bilgilerim", "Şifre Deđişikliği", "API Key", and "Paket Deđişikliği". The "API Key" item is highlighted. On the right, there is a section titled "API Key" with a text input field containing a long string of 'x' characters. Below the input field, there is a "Yenile" (Refresh) button.

Alınan token ile “curl” kullanarak istek gönderimi yapılır. İstek gönderimi, çıktı başlıkları ve çıktı örnekleri için metin sınıflandırma api dökümanı incelenebilir.