

Yapay Zekanın GSB Yazılım Projelerine Entegrasyonu Hackathonu

Proje Takımı: ŞÖBİYET

Takım Üyeleri: Emine Sude ASLAN, Ayşenur YÖRÜR, Eda USLU,
Elif ARSLAN, Ayşe Gül Kılıç

Proje Adı: Kişiselleştirilmiş Spor Öneri

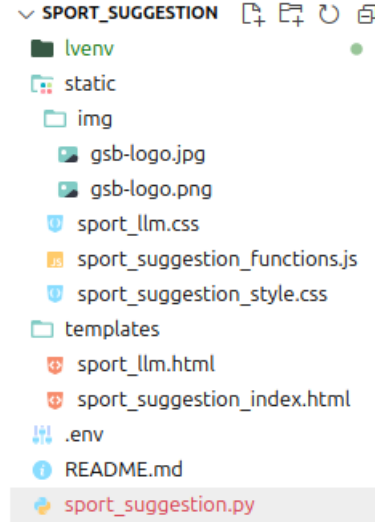
Proje Hedef Kitlesi: Biz Uygulaması ve Web Sitesi Kullanıcıları - Gençler

Proje Amacı: Kullanıcıların spora başlamak için ilk adım olan karar verme aşamasını geçmelerini sağlamak

Proje Konusu: Kullanıcılara belli bilgilerine göre yapay zeka destekli spor önerisinde bulunulması ve kullanıcının büyük dil modeli ile konuşması.

Proje Bileşenleri

Bu kısımda projenin oluşturulmasında kullanılan teknolojiler açıklanmış ve uygulamanın kullanıcı tarafından görünümü gösterilmiştir. Projenin genel dosya yapısı *Şekil 1*'de gösterilmiştir.



Şekil 1: Proje Genel Yapısı

HTML / CSS ve JavaScript: Kullanıcı arayüzünün oluşturulmasında kullanılmıştır. Kullanıcının gayet net şekilde anlayabileceği bir tasarım yapılmıştır. Bunun yanında kullanıcının dikkatini çekebilmesi için canlı bir yeşil tonu (greenyellow) tercih edilmiştir.

Spor Öneri

1. Cinsiyetiniz: ☐ Kadın ☐ Erkek

2. Yaşınız:

3. Boyunuz:

4. Kilonuz:

Gönder

Şekil 2: Uygulama Form Sayfası

Form ve sohbet sayfalarındaki buton fonksiyonlarının oluşturulması için ayrı bir javascript sayfası oluşturuldu (sport_suggestion_functions.js). Javascript kodlarının html uzantılı sayfalarının içine eklenmemesinin en büyük sebebi proje büyüdüğü zaman geliştirici tarafından okunabilirliğini kaybetmemektir.

Flask: Uygulamanın back-end kısmı bir python kütüphanesi olan flask ile hazırlandı. Flask kullanılmasındaki temel sebep anlaşılabilirliğinin daha kolay olması ve python dilinin günümüzde oldukça önemli bir yere sahip olmasından kaynaklanmaktadır. Uygulama dosyasında (sport_suggestion.py) sohbet sayfasında kullanılan dil modeli API'sinin çağırımı yapılmış ve aynı zamanda kullanıcının sohbet ederken spor konusunun dışına çıkmaması için statik bir anahtar kelime kümesi oluşturularak konuların sınırının spor içinde kalması sağlanmıştır.

Yoga veya Pilates

Bana bir şeyler anlatır mısın?

Bu sohbet sadece spor hakkında konuşmak içindir. Lütfen sporla ilgili bir soru sorun!

Yoga neye iyi gelir?

Yoga'nın Faydaları:

Esnekliği ve hareket aralığını artırır:
Yoga pozları, vücudu esnetmeye ve hareket aralığını genişletmeye yardımcı olur.

Gücü ve dayanıklılığı geliştirir:
Karmaşık yoga pozları, kasları güçlendirmeye ve dayanıklılığı artırmaya yardımcı olur.

Duruşu düzeltir:
Yoga, omurga hizalamasını iyileştirerek duruşu düzeltir ve sırt ağrılarını azaltır.

Stresi azaltır ve rahatlamayı destekler:

Sorunuzu yazın...

Gönder

Şekil 3: Dil Modelinin Konu Sınırlamasına Örnek Bir Konuşma

Gemini model API'sinin projeye entegrasyonu açık bir şekilde yapılmak yerine .env isimli bir dosyaya kaydedilmiş ve bu sayede doğrudan erişimin önüne geçilerek güvenliği sağlanmıştır. Python dosyasında "os" kütüphanesi komutları kullanılarak izin işlemleri ile çağırımı yapılmaktadır.

Aşağıdaki açıklamalar hackathon bilgilendirme dosyasının "Değerlendirme Kriterleri" başlığı altında toplanan maddelerin ilk dördüne açıklık getirmektedir.

1. Yapay Zeka Entegrasyonu:

1.1. Model Geliştirme:

Bu kısımda kullanıcının girdiği bilgiler doğrultusunda öneri yapacak olan modelin eğitim süreçleri açıklanmıştır.

Makine öğrenmesi sürecinde klasik adımları takip ederek kendi modelimizi geliştirdik. İlk olarak **problem tanımı** yaparak çözmek istediğimiz konuyu belirledik. Ardından, **veri toplama** aşamasında ilgili verileri derledik ve modelimizin gereksinimlerine uygun hale getirdik.

Sonrasında, **veri ön işleme** sürecinde eksik verileri tamamlayarak, gereksiz değişkenleri temizleyerek ve verileri uygun formata getirerek modelleme için hazır hale getirdik. **Keşifsel veri analizi (EDA)** adımı ise veriyi daha iyi anlamak adına istatistiksel özetler ve görselleştirmeler gerçekleştirdik.

Bu hazırlıkların ardından, **modelleme** aşamasına geçerek uygun makine öğrenmesi algoritmalarını seçtik ve modelimizi eğittik. Modelimizin performansını ölçmek amacıyla **değerlendirme** adımı farklı metrikler kullanarak sonuçları analiz ettik.

Son olarak, elde ettiğimiz çıktıları değerlendirerek **sonuç** kısmında modelimizin başarısını yorumladık ve elde edilen bilgileri raporladık. Böylece, veri odaklı bir yaklaşımla problemimize çözüm sunan başarılı bir model geliştirmiş olduk.

1.1.1 Problem Tanımı

Bu çalışmada, kullanıcıların temel fiziksel özellikleri (boy, kilo, yaş gibi) kullanılarak gençlik spor bünyesinde başarılı olabilecekleri spor branşlarını ve bu branşların olimpiyatlardaki karşılıklarını belirlemeyi amaçlıyoruz. Problem, sınıflandırma tabanlı bir makine öğrenmesi problemi olup, bireylerin fiziksel yeterliliklerine göre en uygun spor dallarına yönlendirilmesini hedeflemektedir.

1.1.2 Veri Toplama

Bu çalışmada, Atina 1896'dan Rio 2016'ya kadar olan 120 yıllık tüm Olimpiyat müsabakalarına katılan sporcuların verileri kullanılmıştır. Veri seti, yaz ve kış oyunlarında yarışan sporcuların fiziksel ve performans özelliklerini içermektedir.

Önemli bir nokta olarak, 1992'ye kadar kış ve yaz oyunlarının aynı yılda düzenlendiği, ancak 1994 yılından itibaren kış oyunlarının farklı bir dört yıllık döngüye geçtiği belirtilmelidir. Bu nedenle, 1994'ten sonra kış ve yaz oyunları farklı yıllarda gerçekleşmiştir (örneğin, 1994 Kış, 1996 Yaz, 1998 Kış vb.).

Veri seti, kaç sporcu, kaç farklı spor dalı ve kaç farklı ülke temsil edildiği gibi temel bilgileri içerirken, aynı zamanda sporcuların en çok hangi ülkelerden geldiği, kimlerin madalya kazandığı ve sporcuların fiziksel özellikleri (cinsiyet, boy, kilo vb.) gibi çeşitli analizlere olanak tanımaktadır.

Bunun yanı sıra, veri setinde Nazi Almanyası'nın ev sahipliği yaptığı yıllarda rakiplerinin üzerinde üstünlük kurduğu veya geçmişte şiir ve resim gibi sanat dallarının da Olimpiyat yarışmalarında yer aldığı gibi ilginç tarihsel bilgiler de bulunmaktadır.

Veri Kümesi Detayları:

Veri seti athlete_events.csv dosyasında yer almakta olup, 271.116 satır ve 15 sütundan oluşmaktadır. Her satır, bireysel bir sporcu ve onun belirli bir yarışmaya katılımını temsil etmektedir.

Veri setindeki sütunlar şunlardır:

- **ID:** Her sporcu için benzersiz kimlik numarası
- **Name:** Sporcu adı
- **Sex:** Cinsiyet (M/F)
- **Age:** Yaş
- **Height:** Boy (cm cinsinden)
- **Weight:** Kilo (kg cinsinden)
- **Team:** Takım adı
- **NOC:** Ülke kodu (Ulusal Olimpiyat Komitesi 3 harfli kodu)
- **Games:** Yıl ve sezon bilgisi
- **Year:** Yıl (tam sayı olarak)
- **Season:** Yaz veya Kış Olimpiyatları
- **City:** Ev sahibi şehir
- **Sport:** Spor dalı
- **Event:** Yarışma kategorisi
- **Medal:** Kazanılan madalya türü (Altın, Gümüş, Bronz veya NA)

Bu çalışma kapsamında, internetten hazır olarak bulunan Olimpiyat sporcularına ait veri seti kullanılmıştır. Veri kümesi, ikinci veri toplama yöntemi ile elde edilmiştir; yani veriler doğrudan toplanmamış, daha önce başkaları tarafından derlenmiş mevcut bir veri seti kullanılmıştır. Bu nedenle, projede ek bir veri toplama süreci gerçekleştirilmemiştir. Veri setinin ham halinde 271116 örnek vardır bu örnekler 15 öznitelikten meydana gelmiştir.

1.1.3 Veri Ön İşleme

Bu çalışmada, Olimpiyat sporcuları veri seti üzerinde veri temizleme ve düzenleme adımları gerçekleştirilmiştir. Veri setinin daha tutarlı ve analiz edilebilir hale getirilmesi için eksik verilerin giderilmesi, gereksiz değişkenlerin kaldırılması, veri tutarsızlıklarının düzeltilmesi ve sınıflandırma için gerekli dönüşümlerin yapılması gibi çeşitli veri ön işleme teknikleri uygulanmıştır. Aşağıda, veri ön işleme sürecinde gerçekleştirilen adımlar ayrıntılı bir şekilde açıklanmıştır. Bu adımların notebook üzerinde her bölümün altında iç görüler olacak şekilde yazılmıştır ve hatalar tespit edilerek sonuçlar elde edilmiştir.

1.1.3.1. Eksik Verilerin Temizlenmesi

Veri setindeki eksik değerlerin giderilmesi için sistematik bir yaklaşım izlenmiştir. İlk olarak, madalya bilgisi dışındaki tüm sütunlarda eksik veri içeren satırlar silinmiştir. Ancak, madalya kazanamayan sporcuların bilgileri korunarak, eksik madalya değerleri "No Medal" olarak atanmıştır. Bu sayede, veri kaybı en aza indirilmiş ve madalya kazanmayan sporcuların da analizlere dahil edilmesi sağlanmıştır.

1.1.3.2. Gereksiz Sütunların Kaldırılması

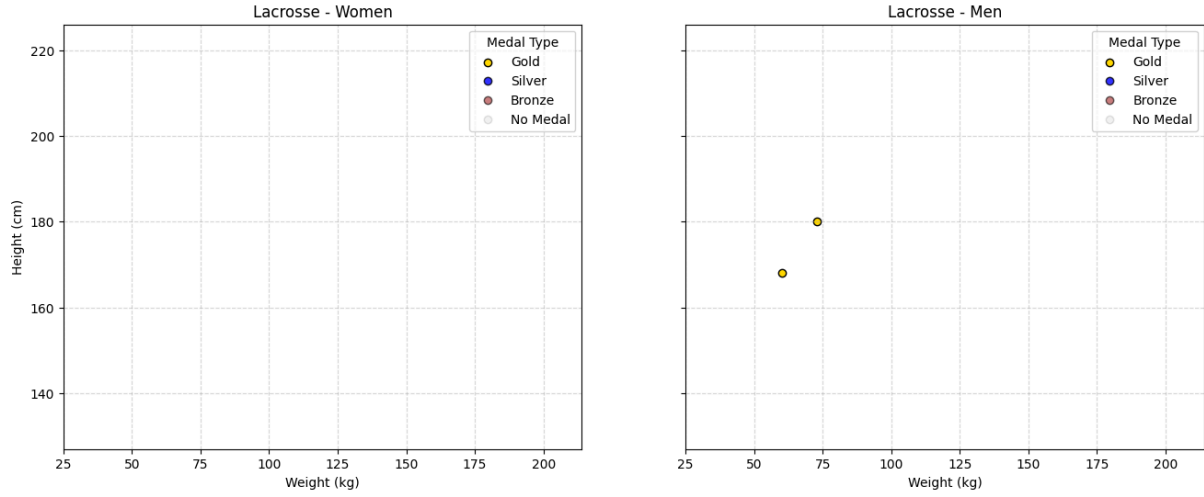
Veri setinde analiz için doğrudan gerekli olmayan sütunlar belirlenerek kaldırılmıştır. Örneğin:

- Takım adı (Team) sütunu, bireysel sporcuları temel alan analiz için gereksiz görüldüğünden çıkarılmıştır.
- Oyunlar (Games), Sezon (Season) ve Ev Sahibi Şehir (City) sütunları, yıl bilgisi zaten ayrı olarak bulunduğu için fazladan bilgi içermekteydi. Bu nedenle kaldırılmıştır.

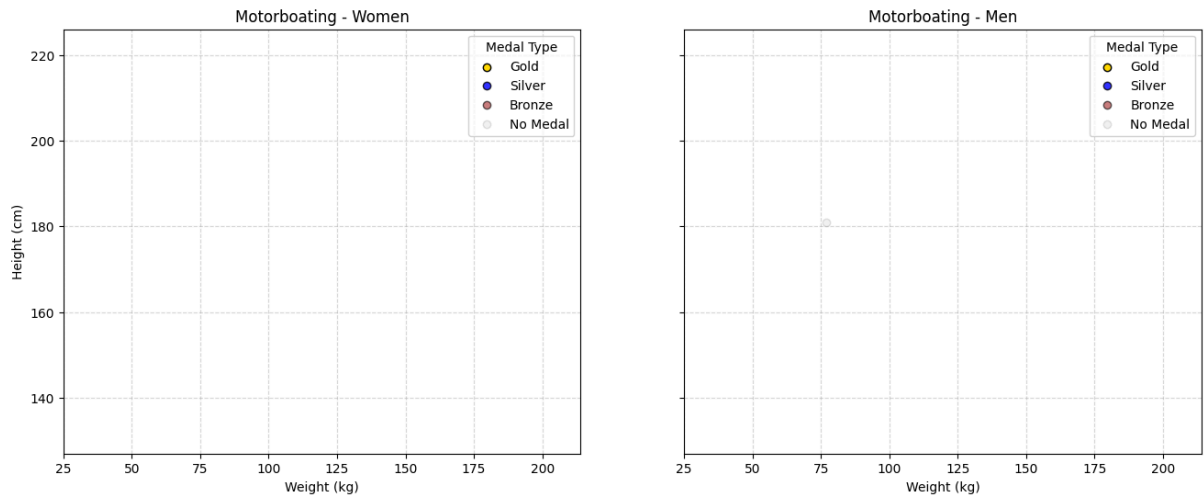
Bu adım, veri setinin gereksiz bilgi yükünden arındırılmasını ve analiz sürecinin daha verimli hale getirilmesini sağlamıştır.

1.1.3.4. Belirli Sporların Çıkarılması

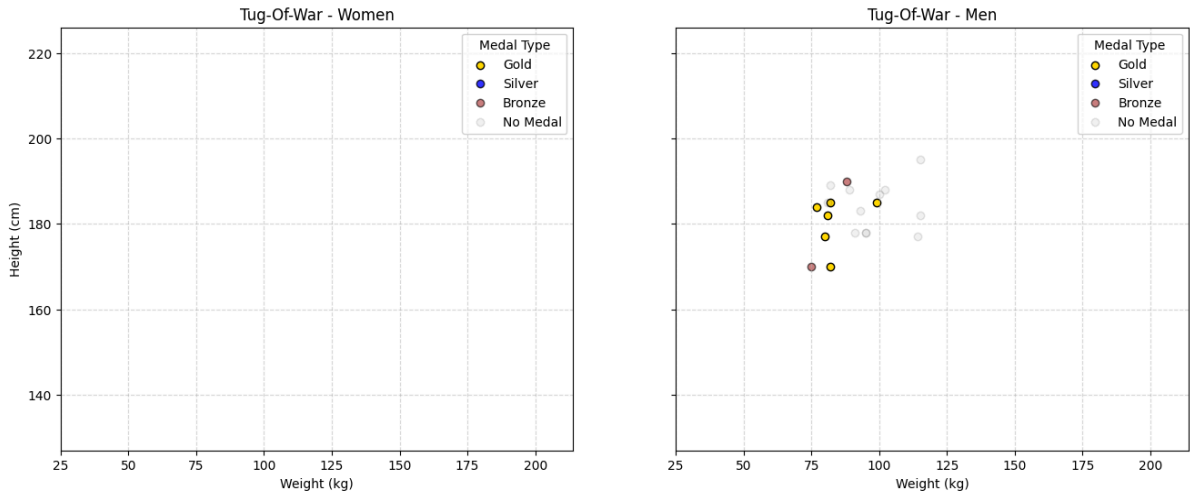
Tarihsel olarak Olimpiyat Oyunlarında yer almış ancak modern analiz kapsamında ilgisiz veya tutarsız görülen bazı sporlar veri setinden çıkarılmıştır. Motorbot Yarışı, Halat Çekme, lakros gibi bazı sporlarda şekilde görüldüğü gibi yeterli örnek yoktu ve bunlar veri seti içerisinden çıkarıldılar. Şekil 4, Şekil 5 ve Şekil 6'da görüntülendiği gibi çıkarılmıştır



Şekil 4. Lacrosse Kadın/Erkek için Boy-Kilo Tablosu



Şekil 5. Motorboating Kadın/Erkek için Boy-Kilo Tablosu



Şekil 6. Tug-of-War Kadın/erkek için Boy-Kilo Tablosu

1.1.3.5. Yinelenen verilerin temizlenmesi

Veri setinde aynı sporcunun birden fazla kez tekrarlanmasını önlemek için yinelenen satırlar tespit edilerek kaldırılmıştır. Bunun için, sporcuların adı, boy ve kilo bilgileri temel alınarak

tekrar eden kayıtlar belirlenmiş ve yalnızca ilk görülen kayıt korunarak diğerleri veri setinden çıkarılmıştır. Bu işlem, analizde yanlış yönlendirici tekrarların önüne geçerek verinin bütünlüğünü artırmıştır.

1.1.3.6. Düşük Katılımcılı Etkinliklerin (Eventlerin) Çıkarılması

Veri setinde spor dalları (Sport) ve bu dalların altındaki spesifik etkinlikler (Event) arasında bir ayrım bulunmaktadır. Bir spor dalına toplamda yüksek sayıda sporcu katılmış olabilir, ancak bu dalın içindeki bazı spesifik etkinlikler zaman içinde düşük katılımcı sayısına sahip olabilir. Bu durum, istatistiksel analiz ve modelleme açısından güvenilir olmayan veri noktaları oluşturabilir.

Bu nedenle, belirli bir etkinliğe (Event) katılan sporcu sayısı zaman içinde 100'ün altında kaldığında, bu etkinlik veri setinden çıkarılmıştır. Örneğin:

- Atletizm (Sport) dalına yıllar içinde toplam 20.000 sporcu katılmış olabilir.
- Ancak Atletizm altındaki "20 kilometre yürüyüş" (Event) etkinliğine toplamda yalnızca 100'den az sporcu katılmışsa, bu etkinlik veri setinden çıkarılmıştır.

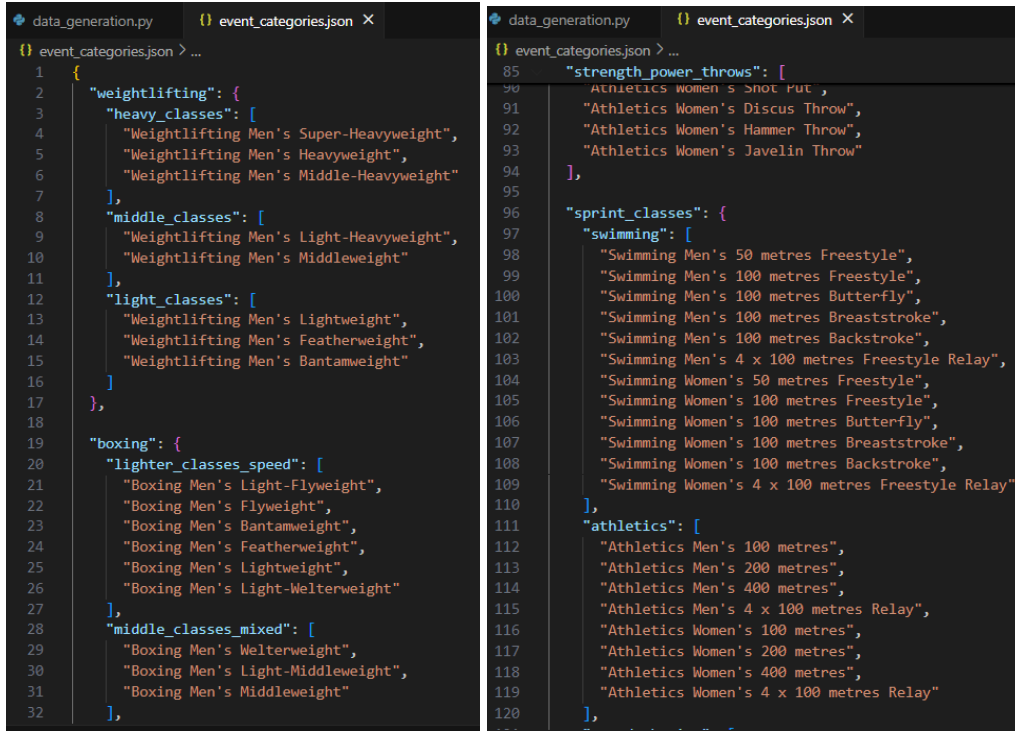
Bu işlem, istatistiksel anlamlılık açısından yeterli veriye sahip olmayan etkinliklerin analiz ve modelleme sürecini olumsuz etkilemesini önlemek amacıyla gerçekleştirilmiştir. Böylece, yalnızca yeterli katılımcıya sahip olan ve istatistiksel olarak anlamlı çıkarımlar yapmaya elverişli etkinlikler veri setinde tutulmuştur.

1.1.3.7. Spor Dallarının Kategorilere Ayrılması

Bu çalışmada, spor dalları ve etkinlikler birleştirilerek yeni bir özellik olan "Kategori" (Feature) oluşturulmuştur. Sporların doğasından kaynaklanan farklı fiziksel ve teknik gereksinimleri göz önünde bulundurularak, bu etkinlikler belirli kategoriler altında gruplandırılmıştır.

Bilindiği gibi, atletizm gibi geniş kapsamlı bir spor dalında kısa mesafe koşuları, orta mesafe koşuları ve maraton gibi farklı branşlar bulunmaktadır. Bu branşlar, sporculardan farklı fiziksel ve performans gereklilikleri beklemektedir. Aynı şekilde, boks ve halter gibi sporlar da farklı ağırlık kategorilerine ayrılmıştır. Bu ayrımı daha iyi analiz edebilmek için, etkinlikler (Event) fiziksel gereksinimler ve performans kriterleri temelinde kategorilere ayrılmıştır.

Bu kategoriler, event_categories.json dosyasında yer almakta olup, spor dallarının fiziksel gereksinimleri, bireysel veya takım sporu olup olmaması gibi çeşitli kriterler dikkate alınarak oluşturulmuştur. Şekil 7'de bu gruplandırmaların bir kısmına örnek verilmiştir.



```
{
  "weightlifting": {
    "heavy_classes": [
      "Weightlifting Men's Super-Heavyweight",
      "Weightlifting Men's Heavyweight",
      "Weightlifting Men's Middle-Heavyweight"
    ],
    "middle_classes": [
      "Weightlifting Men's Light-Heavyweight",
      "Weightlifting Men's Middleweight"
    ],
    "light_classes": [
      "Weightlifting Men's Lightweight",
      "Weightlifting Men's Featherweight",
      "Weightlifting Men's Bantamweight"
    ]
  },
  "boxing": {
    "lighter_classes_speed": [
      "Boxing Men's Light-Flyweight",
      "Boxing Men's Flyweight",
      "Boxing Men's Bantamweight",
      "Boxing Men's Featherweight",
      "Boxing Men's Lightweight",
      "Boxing Men's Light-Welterweight"
    ],
    "middle_classes_mixed": [
      "Boxing Men's Welterweight",
      "Boxing Men's Light-Middleweight",
      "Boxing Men's Middleweight"
    ]
  },
  "strength_power_throws": [
    "Athletics women's Shot put",
    "Athletics Women's Discus Throw",
    "Athletics Women's Hammer Throw",
    "Athletics Women's Javelin Throw"
  ],
  "sprint_classes": {
    "swimming": [
      "Swimming Men's 50 metres Freestyle",
      "Swimming Men's 100 metres Freestyle",
      "Swimming Men's 100 metres Butterfly",
      "Swimming Men's 100 metres Breaststroke",
      "Swimming Men's 100 metres Backstroke",
      "Swimming Men's 4 x 100 metres Freestyle Relay",
      "Swimming Women's 50 metres Freestyle",
      "Swimming Women's 100 metres Freestyle",
      "Swimming Women's 100 metres Butterfly",
      "Swimming Women's 100 metres Breaststroke",
      "Swimming Women's 100 metres Backstroke",
      "Swimming Women's 4 x 100 metres Freestyle Relay"
    ],
    "athletics": [
      "Athletics Men's 100 metres",
      "Athletics Men's 200 metres",
      "Athletics Men's 400 metres",
      "Athletics Men's 4 x 100 metres Relay",
      "Athletics Women's 100 metres",
      "Athletics Women's 200 metres",
      "Athletics Women's 400 metres",
      "Athletics Women's 4 x 100 metres Relay"
    ]
  }
}
```

Şekil 7. Json kateogrileri

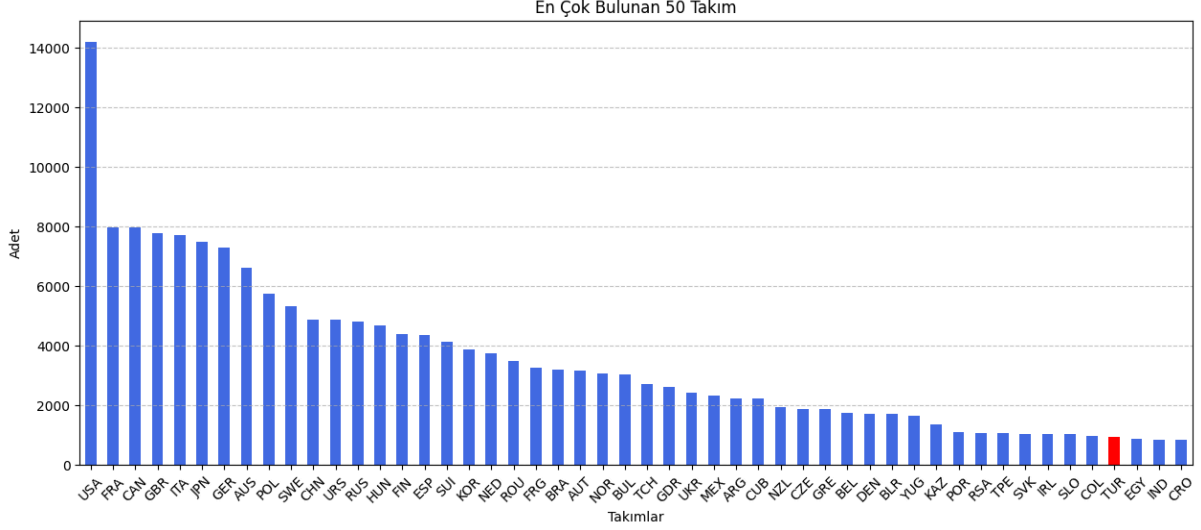
Bu gruplandırmalar tamamen içsel değerlendirme ve mantıksal çıkarımlar ile gerçekleştirilmiştir. Sporların gerektirdiği fiziksel yetkinlikler (güç, hız, dayanıklılık vb.), bireysel veya takım bazlı olmaları, teknik beceri gereksinimleri ve yarış formatları gibi unsurlar dikkate alınarak oluşturulmuştur.

Ayrıca, bu gruplandırma ileriye dönük analiz ve model geliştirme süreçlerinde daha anlamlı öngörüler elde etmeyi sağlayacak bir temel oluşturmuştur. Kategorilendirme sayesinde, sporcuların belirli etkinliklerde başarı ihtimalini analiz etmek ve fiziksel uygunluklarına göre en doğru spor dalına yönlendirilmesini sağlamak amaçlanmaktadır.

1.1.4. Keşifsel Veri Analizi

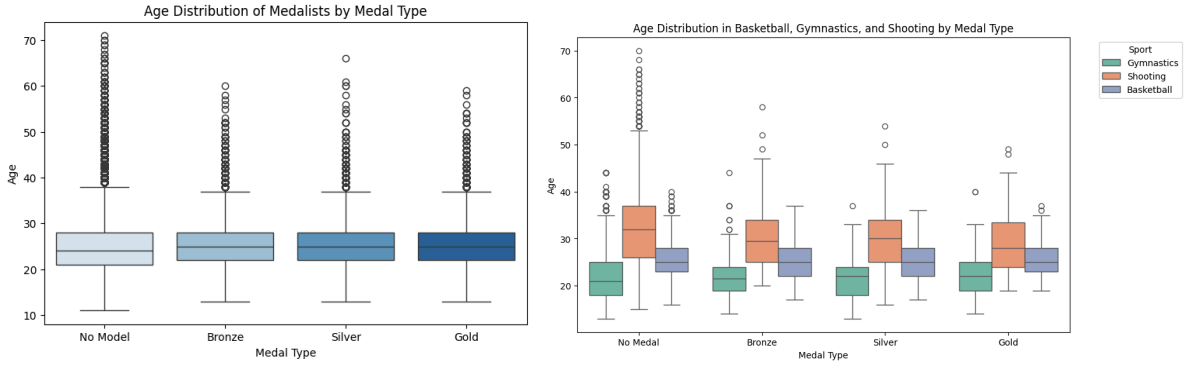
Takım isimleri, sporcuların fiziksel özellikleri veya doğdukları ülke hakkında doğrudan bilgi sağlamasa da, ülkelerin belirli spor dallarına verdikleri önemi ve yaptıkları yatırımları yansıtabilir. Özellikle, farklı ülkelerden gelen sporcuların belirli bir takımda yarışabilmesi nedeniyle, takım ismi bireysel analizler için yanıltıcı olabilir. Ancak, ülkelerin belirli spor dallarında nasıl bir rekabet içinde olduğunu görmek açısından önemli bir değişken olduğundan takım bilgisinin veri setinde tutulması planlanmaktadır.

Şekil 8'de görüldüğü gibi, en fazla yarışmalara katılan takımlar, ülkelerin spor alanındaki yatırımlarını ve Olimpiyatlar'daki aktif rollerini göstermektedir.



Şekil 8. En çok olimpiyatlara katılan ülke

Şekil 9'da incelediğimizde, madalya kazanan sporcuların ağırlıklı olarak 20'li yaş grubunda yer aldığı görülmektedir. Ancak, bazı outlier sporcular, özellikle bedensel olarak yoğun efor gerektirmeyen branşlarda madalya kazanmış olabilir. Örneğin, silahlı atıcılık gibi fiziksel dayanıklılıktan ziyade odak ve teknik beceri gerektiren spor dallarında yaşça büyük sporcuların madalya kazanma olasılığı daha yüksek olabilir. Bu durumu daha detaylı incelemek adına, spor bazında madalya kazanan yaş gruplarını Şekil 9'da analiz ettik.



Şekil 9. Yaş ve Sporların Madalya kazanımı ile ilişkileri

Projede kullanıcı ilk sayfada yaş, cinsiyet, boy ve kilo bilgilerini girdiği bir formla karşılaşılıyor. Bu formu doldurduktan sonra bir sohbet sayfasına yönlendiriliyor. Sohbet sayfasında kullanıcı Gemini-Pro dil modelinin entegre edildiği bir arayüzle konuşmaktadır.. Kullanıcının spor dışında bir konu konuşarak amacından çıkmasının önüne geçebilmek için kısıtlamalar uygulanmıştır.

1.1.5 Modelleme

Öncelikle, ek özellik mühendisliği (feature engineering) veya kategorileştirme yapılmadan bir base model oluşturulmuştur. Bu aşamada, farklı makine öğrenmesi algoritmaları (Random Forest, Logistic Regression, Support Vector Machine, Gradient Boosting, Naive Bayes) denenerek spor tahmini yapılmıştır. Ardından, modellerin performans sonuçları karşılaştırılmış ve ilerleyen değerlendirmeler sonucunda

1.1.6. Değerlendirme

Bu sonuçlara göre Gradient Boosting en yüksek doğruluk ve F1 skorunu verse de, hesaplama süresi açısından daha verimli olan Random Forest seçilmiştir. Kısıtlı zaman nedeniyle hiperparametre optimizasyonu yapılamamış ve en uygun hız-performans dengesi sağladığı için Random Forest tercih edilmiştir.

Daha sonraki aşamada, kategori bazlı feature mühendisliği eklenerek modelin öneri kapasitesi artırılmış ve 10 farklı event önerisi yapabilen bir model geliştirilmiştir. Bu iyileştirilmiş modelin Top-10 Accuracy sonucu 0.6706 olarak elde edilmiştir.

Model (Base)	Accuracy	F1
Logistic Regression	0.2283	0.1306
Support Vector Machine (SVC)	0.2207	0.1070
Gradient Boosting	0.3256	0.2689
Naive Bayes	0.0702	0.0357
Random Forest	0.2627	0.2327

Model (with new features)	Accuracy
Random Forest	0.6706

1.1.7 Sonuç

Geliştirilen kategorik 10 tahmin yöntemi ile kişiye 60 spor dalı arasından 10 öneri sunulmaktadır. Böylece bireylerin kişilik özellikleri ve tercihlerine göre seçim yapma şansı sağlanmaktadır. Gelecek planlarımız arasında LLM (Large Language Model) entegrasyonu bulunmaktadır. Bu sistemde, gençlere kişilik analizi yapıldıktan sonra, makine öğrenmesiyle belirlenen 10 spor dalı içinden en uygun 3 dal önerilecektir. Örneğin, yüzme ile ilgilenen bir bireyin fiziksel yetenekleri yüzmeye yatkınsa, yüzme önerilecektir. Örneğin, Şekil 10'da önerilen 2. tahmin gerçekte olimpiyat sporcusunun yaptığına girmektedir

```
[52] top_10_labels[0]
array(['team_sports - football',
      'athletics_technical - hurdles_steeplechase', 'fencing',
      'endurance_stamina - triathlon',
      'athletics_technical - jumps_and_vault',
      'watercraft_racing - rowing',
      'endurance_stamina - cross_country_skiing',
      'sprint_classes - swimming', 'endurance_stamina - running',
      'sprint_classes - athletics'], dtype=object)

y_valid.iloc[0]
'athletics_technical - hurdles_steeplechase'
```

Şekil 10. Tahmin ve Validation

2. Kullanılabilirlik ve Dökümantasyon:

Kullanıcı uygulamayı açtığı anda ,Biz uygulamasındaki/ web sitesindeki butona tıkladığı an da diyebiliriz, karşısına bir form çıkıyor. Form herkes tarafından kolayca anlaşılabilmesi için oldukça sade bir yapıda tasarlanmıştır. Formu doldurduktan sonra kullanıcının yönlendirildiği sohbet ekranı da aynı tasarım göz önünde bulundurularak tasarlanmıştır ve bu sayede sayfa değişmiş olmasına rağmen anda kalmaları hedeflenmiştir. Bu kısım ile ilgili kullanıcı herhangi bir dökümana ihtiyaç duymamaktadır. Geliştirici tarafında ise yorum satırları ile okunabilirlik kolaylaştırılmıştır.

Model geliştirme kısmında notebook içerisinde diğer geliştiriciler için iç görülerimizi barındıran metinler mevcuttur. Aynı zamanda daha sonrası için geliştirdiğimiz modelin uygulamaya entegrasyonu için scriptleme ve docker'a geçilmiştir ama tamamlanamamıştır.

3. Proje İşlevselliği ve Teknik Uygulama:

Raporun başında da belirttiğimiz gibi asıl hedefimiz olan gençlerin spora başlama yolculuklarında ilk adımlarını atabilmeleri için karar aşamalarına destek olarak bu süreçte aksiyon almalarını kolaylaştırmak amaçlanmıştır. Eğitilen model sayesinde kullanıcının fiziksel özelliklerine uygun olabilecek bir / birkaç öneri alması ve sonrasında spor konusunda konuşabileceği bir dil modeline yönlendirilmesi sağlanarak kullanıcının önerilen spor hakkında odağını kaybetmeden edinmek istediği bilgilere erişmesi hedeflenmiştir.

4. Güvenlik ve Gizlilik:

API'nin Korunması: Dil modelinin kullanılması için kullanılan API değeri doğrudan koda yazılmak yerine dizinde başka bir dosyanın içine kaydedilmiştir.

Minimal Veri Politikası: Projede kullanıcıya dair sadece ihtiyaç duyulan bilgiler istenmektedir.

Anonimlik: Kullanıcıya ait kimlik numarası, isim - soyisim gibi bilgiler istenmeyerek anonimlik sağlanmıştır.

Verilerin Saklanmaması: Kullanıcının girdiği veriler ve dil modeli ile yapılan konuşmalar herhangi bir şekilde tutulmayarak verilerin tekrar erişilmesinin önüne geçilmesi sağlanmıştır.

Biz uygulamasının Güvenlik Önlemleri: Biz uygulamasına entegre edilmesi hedeflenen Spor Öneri projesi sistemin halihazırda var olan güvenliğinden faydalanması düşünülmüştür. Web adresinin "https" protokolü ile iletişim sağlıyor olması ağda dolaşan kişilerin bilgilerinin güvenli şekilde iletiliğinin işaretidir.

GELİŞTİRME ÖNERİLERİ

Bu kısımda tarafımızdan hazırlanan projenin ilerleyen evrelerinde eklenebilecek çeşitli geliştirme önerilerinden bahsedilmiştir.

- Sistemimiz, **kullanıcının konumu ve önerilen spor dalı** doğrultusunda **en uygun gençlik merkezlerini** de sunarak daha kişiselleştirilmiş bir deneyim sağlamayı amaçlamaktadır. Örneğin, **Adana’da yaşayan bir kullanıcıya model karate sporunu önerdiğinde, Adana’daki karate branşı bulunan gençlik merkezleri** de kullanıcıya bildirilecektir.
- Şu an sistem, **kategori bazlı 10 farklı spor dalı önerisi** sunmaktadır. Ancak, ilerleyen aşamada LLM (Large Language Model) entegrasyonu ile kullanıcıya **spora yönelik kişilik analizi** yaparak önerileri daha da kişiselleştirmeyi hedefliyoruz. Böylece, bireyin **fiziksel uygunluğu** kadar **zihinsel ilgisini ve motivasyonunu** da göz önünde bulundurarak daha **doğru ve tatmin edici** spor önerileri sunulabilecektir.