1. **Amaç:**

Bu çalışmanın amacı EEM612 kodlu Örüntü Tanıma ve Makine Öğrenmesi dersi kapsamında dönem ödevinde yapılacak çalışmanın ders öğretmeni Dr. Selda GÜNEY hn. ile birlikte belirlenmesi çalışmalarını dokümante etmektir.

1. **Alternatifler:**
   1. **Alternatif-1:**

Li-Ion BMS

<https://uk.mathworks.com/help/stats/deploy-neural-network-regression-model-to-fpga-platform.html>

<https://data.mendeley.com/datasets/cp3473x7xv/3>

<https://www.sciencedirect.com/science/article/pii/B9780444595133000157?via%3Dihub>

**SE Yorum:** Artı ve eksileri aşağıda sıralanmıştır.

Artılar:

* HW implementasyonu olması sebebiyle ilgi çekici. Şuan için MATLAB yerine Python ile veri seti içerisindeki bazı veriler üzerine ML algoritmaları hazırlayabilirim. Belki daha sonra Yüks.Lis. bitirme projesi olarak HW (FPGA veya başka bir donanım) ile uygulama düşünülebilir.
* Figes firmasında MATLAB Uygulama Müh. olarak HDL Coder ailesi sorumluluğunda Oca.2019 – Eyl.2020 arasında yaklaşık 2 yıl gibi bir deneyimim var. MATLAB kısmında kolay adapte olup ilerleyebilirim.

Eksiler:

* Bitirme projesini de kapsayacak ise, Li-Ion bataryalar için BMS konusunda orta üzeri bilgi sahibi olmak gerekecek.

Leonardo Türkiye ve şuan çalışmakta olduğum Lentatek firmalarında şarjlı piller ve Li-Ion bataryalar üzerine çalışmalarım oldu. Buna rağmen yukarıdaki 3.üncü linkten görülebileceği gibi BMS konusu kapsamlı ve önemli bir konudur.

Konunun kendimi geliştirmek istediğim bir alan olmaması, ve EEM612 dersi ve özellikle olası Bitirme Projesi kapsamında konseptin detaylı bilinmesinin gerekli olduğunu düşünmem sebepleriyle bu seçeneği seçmek istemiyorum.

Belki sizin farklı bir yorumunuz olur diye bu seçeneği raporlamam gerektiğini düşündüm.

Şöyle ki bu verisetini seçebilirsin ama tezde bunu yapmak zorunda değilsiniz.Algoritmada kazandığın deneyimi tez konunuza da aktarabilirsiniz. O yüzden düşünülebilir.

* 1. **Alternatif-2:**

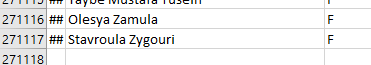
<https://www.kaggle.com/datasets/mysarahmadbhat/120-years-of-olympic-history>

Aslında amacım ML konularını öğrenebileceğim örnek bir çalışma yapabilmek olması sebebiyle, Sisteme Genel Bakış veya Literatür Araştırması diyebileceğimiz konuyla ilgili genel bilgilerin verilmesi kısmını çok da önemsememeyi tercih ediyorum. Böylelikle işime çok da yaramayacak konuları sınırlı zamanımı harcayarak derinlemesine veya en azından tutarlı ve mantıklı bir çerçevede öğrenme çalışmasından kazanç sağlayabileceğimi düşünüyorum.

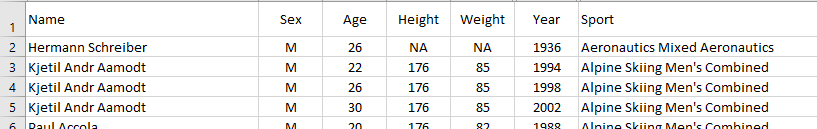
Bu yaklaşım ile

* Yağmur Rejimi gibi veri toplama seçenekleri üzerinde düşündüm.
* 400 – 450 kişi gibi bir yemekhanemiz var. İnsanlara o gün ne kadar iyi kahvaltı yaptıkları, kendilerini aç hissedip hissetmedikleri, boy/kilo/yaş/cinsiyet bilgileri ve “Kendilerine yemekten önce kaliteli ve güzel bir çikolata verilse, tek koşulun bunu yemek zorunda olmaları” teklifine ne yanıt verecekleri sorularak, farklı kişilerin kendi sağlıklarına ne kadar dikkat ettiğiyle ilgili bir veri seti oluşturmayı düşündüm. Çikolota yerine şeker (akide şekeri gibi) ile farklı bir inceleme yapabileceğimiz. Fakat, firmamda 11:30, 11:45, 12:00, 12:15, 12:30, 12:45, 13:00 ve 13:15 şeklinde sekiz (8) grup yemek yiyoruz. Hep aynı sıra ile. Acıkma katsayısı diyebileceğimiz değişkenin etkisinin yüksek olacağını, bunun da bir veri olduğu ve hatta ileriki gruplardan rastgele seçilen bazı kişilere ilk yemek saatinde bu soruya ne cevap verecekleri de sorulabilir gibi seçenekler düşünüldü. Toplamda bu fikir üzerinden gitmeme kararı alındı. Bu kadar kişiyle konuşmak oldukça zaman alıcı bir iş gözüküyor. Ders için olan süreden verileri organize etmekte zorluk çekebilirsiniz.

* Linkten ulaşılabilen veri setiyle amaçlanan ML konseptinin öğrenilme çalışmasının kolaylıkla yapılabileceği kararına varıldı.
* Bazı spor dalları seçilir: Splinter (100m, 200m ve 400m koşucular), Maratoncular, Yüzme, Halter, Dekatlon vb., Yüksek Atlama, Cirit Atma gibi.
* Olimpiyatlara katılan sporcuların fit oldukları varsayılacak.
* Kilo/Boy endekslerine göre (belki Yaş ve Cinsiyeti de dahil ederek) spor dalları için ideal Kilo/Boy endeksleri belirlenecek. Yeni bir olimpiyat sporcusunun Kilo ve Boy bilgileri ile bu sporcunun hangi spor dalında yarıştığı tahmin edilecek. Bu veri setinin oluşturulma amacı ile ilgili aklımda birşey canlanmadı. Neden hangi spor dalında yarıştığını tahmin etmeye çalışıyoruz. Bu açıdan motivasyonu bana biraz zayıf geldi. Eğer farkında olmadığım bir motivasyon kaynağı varsa belirtiniz lütfen.
* SE: 270K satırlık bir veri seti.



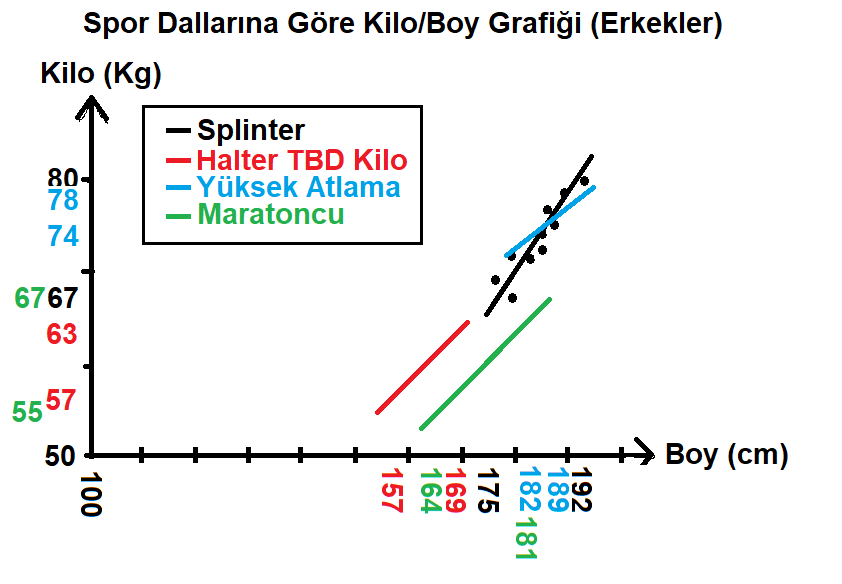
Aşağıdaki sütunlar kullanılabilir.



Cinsiyetlere ve yapılan spor dalına bağlı olarak kilo/boy oranı belirlenecek.

Mesela veriler şöyle olsun:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (cm) | | (Kg) | |
|  | En Kısa | En Uzun | En Hafif | En Şişman |
| Splinter | 175 | 192 | 67 | 80 |
| Halterci | 157 | 169 | 57 | 63 |
| Yüksek Atlama | 182 | 189 | 74 | 78 |
| Maratoncu | 164 | 181 | 55 | 67 |



Şekilde splinterlar için diyelim veriler noktalar ile belirtilenler. Buna bağlı olarak kilo/boy grafiği her spor dalı için belirlenebiliyor. Şekilde farklı renklerle gösterilmiştir.

Grafiklerde min/max değerleri TBD kadar esnetilecek.

Yani elimizdeki verilerle her spor dalı için cinsiyetlere ve yapılan spor dalına bağlı olarak ideal kilo/boy eğrileri olacak.

Not: Olimpiyatlara katılan sporcuların neredeyse eşit fitlikte oldukları varsayılmıştır. Böylece bu verilerle her spor dalı için farklı eğriler elde edebileceğimizi düşünüyorum. Spor ile uğraşmayan ve bu kadar üst seviyede sporcu olmayan kişilerin verileri kullanılmış olsa idi kilo/boy eğrisi anlamlı olmayacak idi.

Resimde görülebileceği gibi splinter eğrisi ile yüksek atlamacı eğrileri bir noktada kesişiyorlar. Ayrıca tahmin algoritması için bu nokta veya birden fazla eğriye eşit uzaklıkta olan noktalar için farklı yaklaşımlar geliştirilmesi gerekecek. Mesela bu nokta için bir grubun olasılığı diğer grup(lar) için daha yüksek olacaktır muhtemelen. Bu tarz bir yaklaşımla veri setimizin verdiği bilgilerle daha doğru tahmin yapacak bir algoritma yapılabilir diye düşünüyorum.

**Motivasyonum:** Ödevin sisteme genel bakış kısmıyla fazla haşır neşir olmadan ML teorisini öğrenmeyi ve bir uygulama ile düşüncelerimi hayata geçirebilme çalışması yapmayı daha değerli buluyorum. İşin hikaye kısmının gerçek verilerle birebir örtüşmesi kaygısı duymayacağım. Hayali bir senaryo için ML uygulaması yapılacak.

**Tez hakkında bir bilgi:**

Bitirme Tezi olarak aşağıdaki seçenek bir alternatif olarak şekillenebilir.

Bunun haricinde TensorFlow Lite veya TinyML gibi IoT’li veya tercihen IoT’siz mikro denetleyici / işlemci veya FPGA çözümleri gibi donanım tabanlı çalışmalar istiyorum.

* İtalya’da bir üniversitede Quantum Comp. Academy takımı ile çalışma yapma şansım var. Bu güzel bir fırsat. Tez konusu için muhakkak düşünülmeli. Ama gömülü yazılımı bu ders kapsamında zaman kısıntından dolayı zorluk çekme ihtimaliniz olur diye düşünüyorum. Eş zamanlı olarak bir iş yerinde çalıştığınız için.

<http://quasar.unina.it/>

[www.qca.unina.it/en](http://www.qca.unina.it/en)

* Önümüzdeki sene açılırsa aşağıdaki dersi almayı düşünüyorum. Bu çalışmalar ile bir tez hazırlayabilirim diye düşünüyorum.

<https://catalog.metu.edu.tr/course.php?prog=567&course_code=5670614>

Li-Ion BMS kısmıyla ilgilenmememin sebebi biraz daha işime yarayacak bir şeyler öğrenme isteği. Kısmet. ☺

Aspilsan BMS’li bataryalar üretebiliyor.

<https://www.aspilsan.com/cozumler/sistem-bataryalari/bb-2590-u/>

Evet, Türkiye’de bile pazarı olan bir alan BMS. Ama diğer iki konunun (gömülü ML ve Quantum Comp.) yanında çok cılız kalıyor.

<https://bilgisayarkavramlari.com/category/kuantum-hesaplama/>

linkinden ulaşılabilecek bir çalışma bile yapsam tez için güzel bir çalışma olacaktır.

Not: Tabi kolay bir tez de seçeneklerim arasında olursa iş/okul dengesi için yararlı olacaktır. Bu düşünceyle siz hocalarım ile tez belirlemek benim için daha faydalı olabilir.

Sizin bir yorumunuz olur mu acaba?

Tez için bir dönem daha vaktiniz var bunu biraz daha kapsamlı bir araştırmayla karar vermekte fayda var. Bol bol son yıllara ait makaleler okuyup karar verilebilir.

* 1. **Alternatif-3:**

<http://www.cs.cmu.edu/~ninamf/courses/401sp18/projects.html>

içerisinde bazı proje önerileri var.

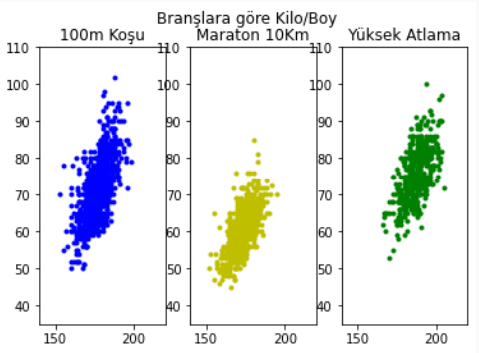
* <https://www.cs.cmu.edu/~tom/pubs/science2008.pdf>

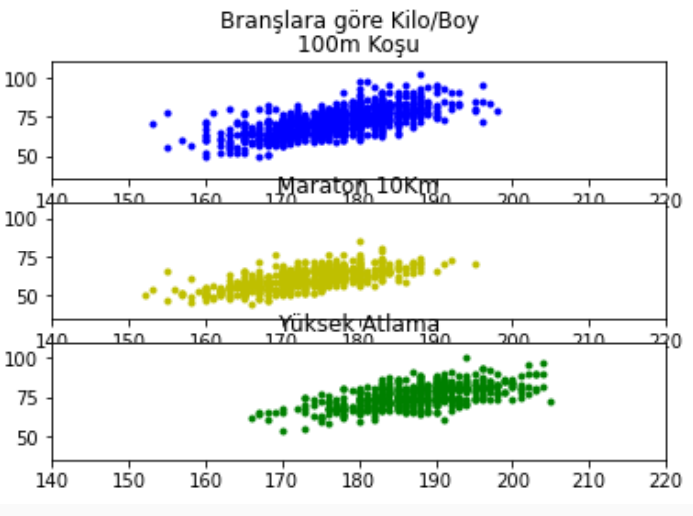
<http://www.cs.cmu.edu/afs/cs/project/theo-73/www/science2008/data.html>

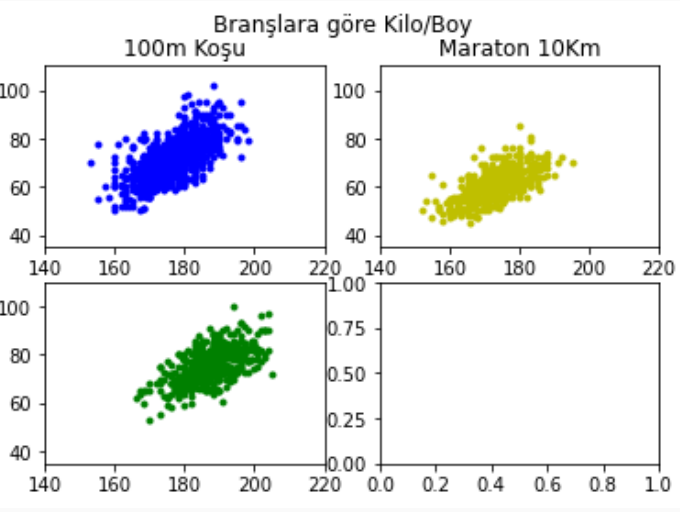
1. **Taslak Çalışmalar**
   1. **Alternatif-2 Üzerinde Çalışma:**

<https://www.kaggle.com/datasets/mysarahmadbhat/120-years-of-olympic-history>

İlgili linkten indirilen veri seti aşağıdaki branşlar için grafik haline getirildi.

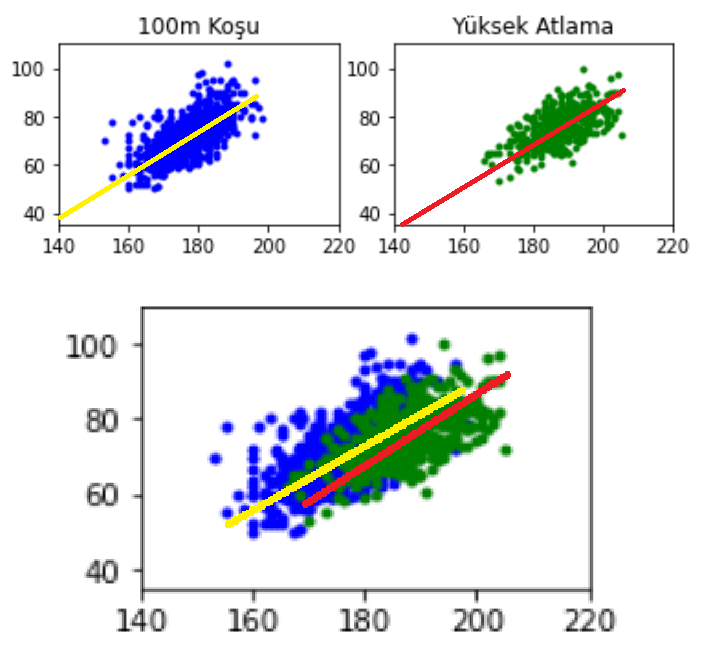


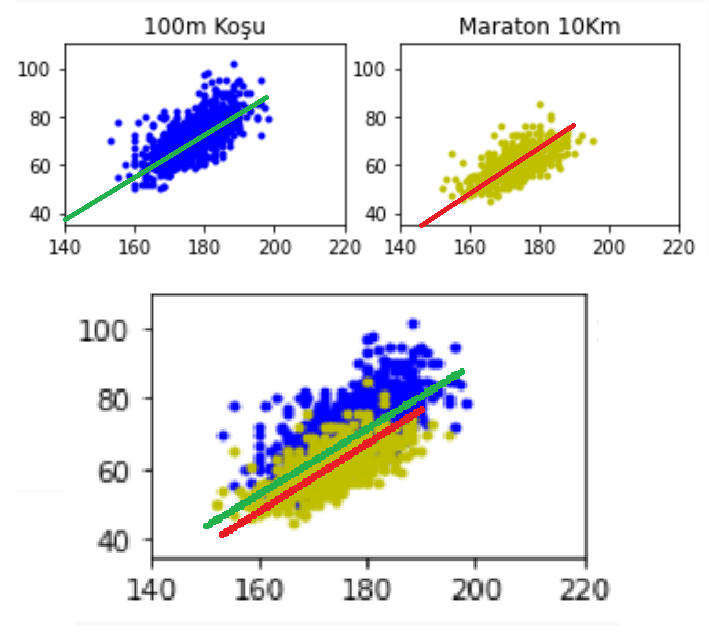


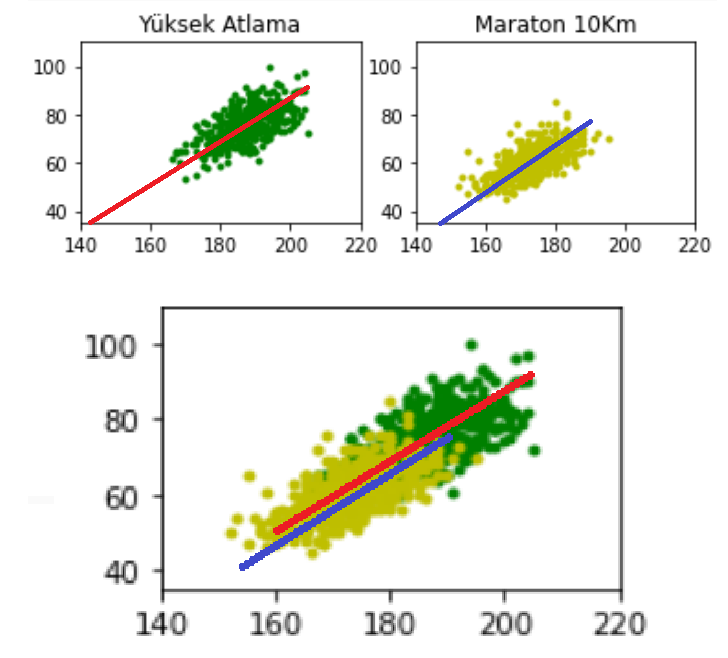


* Athletics Men's 10,000 metres (Maraton)
* Athletics Men's 100 metres (100m Koşu – Erkekler)
* Athletics Men's High Jump (Yüksek Atlama)

Branşlar arasındaki farkı daha net görebilmek için aşağıdaki çalışmalar yapılmıştır.







Toplu gösterim

Sonuçlar:

1. Gözle ve manuel yapılan lineer bir tahmin yaklaşımı uygulandığında bu 3 branşa göre Kilo/Boy grafiklerinin birbirinden farklı ve ayırt edici olacağı görülmüştür.

