

Bu kılavuz, bana birçok kere gelen sorular üzerine hazırlanmıştır. Genç olsun, profesyonel olarak çalışmış olsun, birçok arkadaşım ve öğrencim, birçok yeni mezun, aynı soruları soruyor. Bende bu arkadaşlara yardımcı olmak istedim.

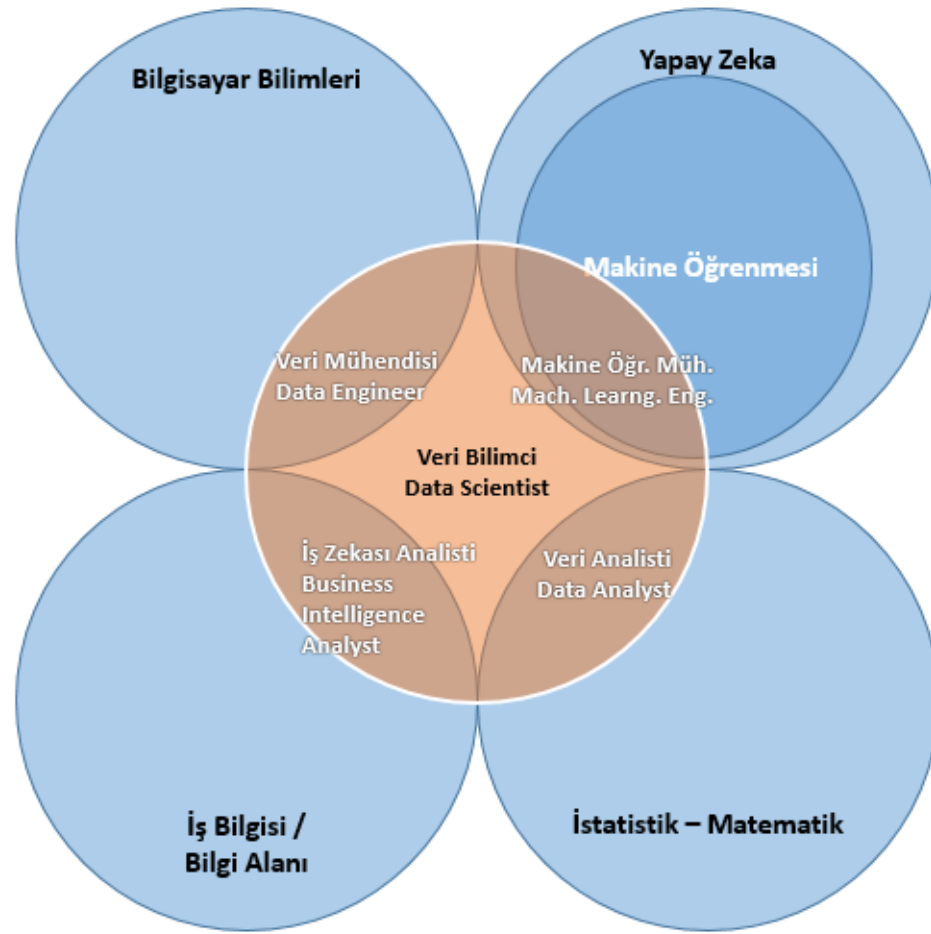
Emre YAZICI
YAPAY ZEKA BİLİM İNSANI

YAPAY ZEKA FAKÜLTESİ
www.yapayzekafakultesi.com.tr

Veri bilimine yeni giriş yapacaksınız, ya da yeni mezun oldunuz veya öğrencisiniz.. Belki de kendinizi geliştireceksiniz. Peki, nasıl yapacaksınız.

NASIL BAŞLARIM: İPUÇLARI

KARAR VERİN	Ne yapmak istediğinizi, neden yapmak istediğinizi iyi düşünün. "Para kazanmak için" düşüncesiyle seçim yapıyorsanız, daha az zahmetle daha çok kazandıran işler var, ticaret gibi.
KENDİNİZİ TANIYIN	Bu kılavuz içinde, karakter testleri ve bu testlerin sonucuna göre becerileriniz doğrultusunda meslek seçimleri haritaları bulunuyor. Önce kendinizi tanıyın ve neyi yapıp neyi yapamayacağınızı belirleyin.
BİLGİNİZİ ÖLÇÜN	Hangi konuda bilgi sahibi değilseniz, iyi tespit yapın, ve nasıl öğreneceğinizi anlatan ipuçlarını takip edin.
ODAKLANIN	Kafanızda, acabalar varsa, yeterince düşünmemişsiniz veya inanmamışsınız demektir.
PROGRAM BELİRLEYİN	Bir program belirleyin ve ona uymaya çalışın. Unutmayın ki, azimli ve sever bir şekilde yapılan işlerde çok hızlı ilerlenir ve başarı elde edilir.
BİR EĞİTİM PROGRAMINA KATILIN	Yapay Zeka Fakültesi, eğitim programına katılın. Başka yerlerde olabilir, ama, YZF'ne garanti veririm
ARAŞTIRIP, DÜŞÜNÜN	Sürekli, "ben bunu nasıl yapardım", "acaba bunu nasıl yapmışlardır" şeklinde düşünün. Özellikle, bir yere gidip gelirken ki boş vaktinizi (trafik vs) değerlendirin.
PROBLEM BULUN	Yer yüzünde icatların hemen hemen birçoğu (teoremler hariç), bir problemi çözmek bir işi kolaylaştırmak için bulunmuştur. Sizde, kolaylaştırılabileceğiniz bir işi veya problemi bulup onu yapay zeka ile nasıl çözerdim diye düşünün.
MENTÖR	Bir mentör veya danışacağınız birisinin olması her zaman iyidir. Düşünsenize, "mentöre sormak istiyorum" joker hakkınız bulunuyor.
NEYİ NEDEN YAPTIĞINIZI BİLİN	Bir siteden 10-20 satır Python kodunu kopyalayıp yapıştırıp çalıştırınca çıkan sonucu linkedin e koyunca hemen "veri bilimci" oldum diye düşünmeyin. Kopyala-Yapıştır kadar basit bir iş yapmış oldunuz. Bu yöntem hiç bir zaman çalışmaz. İki değişkenin korelasyonu yüksek çıktı diye bu değişkenler ilişkili demek değil her zaman !!! Ya da, verilerinizi hadoop a attınız diye şirketiniz yapay zekaya geçmedi !!! Veya Imbalanced bir veri kümesinde %90 accuracy çıkardığınız diye o model başarılı olmadı !!! Neyi neden nasıl yaptığınızı anlarsanız, gerçek model kurabilir, ve sonuç üretebilirsiniz. Unutmayın ki, yapay zeka projelerinin geri dönüşleri bazen çok uzun zaman alabilir, gerçek üretkenlik, "başarı-accuracy" de aldığınız rakam değil, şirkete kazandırdıklarınızdır !!!

**Veri Görselleştiricisi:**

Binlerce hatta milyonlarca satır, onlarca hatta yüzlerce kolon verileri, her tür seviyede (geliştirici, analist, müdür, üst düzey yönetici) sunabilecek halde görselleştiren ve ilgili veri kümesinden, insightlar (önemli bilgiler) çıkaran, bir nevi veri analistine benzer işler yapan uzmandır.

Veri Güvenlik Uzmanı:

Verileri, içlerinde analitik bilgi kalacak ancak kişisel veya kurumsal bilgi kalmayacak şekilde tasarlayan uzman.

Veri Etiketleme Uzmanı:

Derin öğrenme veya makine öğrenmesiyle eğitilecek sistemler için verileri etiketleyen uzman.

Veri Mühendisi: Yandaki grafikte görüldüğü üzere; Veri Mühendisi, veri biliminin, “bilgisayar bilimleri” yönündeki uzantısıdır. Yani genel olarak veri bilimini bilmeli ancak, ağırlıklı olarak, veri hazırlama, veri depolama, veri modelleme, verinin kayıt edilip işleneceği sistemleri modelleme ve tasarlama ile ilgili uzman kişidir. Structured ve özellikle unstructured veri modelleme konusunda, arka zeminde çalışan ve anlık çağrılabilen (web service) ya da sürekli olarak çalışan sistemler konusunda uzman kişidir. İstatistiksel modeller ve yapay zeka konusunda çok detay bilgiye sahip değildirler ancak iş zekası bilgisi konusunda tecrübelidirler.

Big Data Engineer şeklinde de ifade edilebilir. Bugün aslında en çok ihtiyaç olan alanlardan biridir. Ülkemizde, verinin toplanması ve düzenli bir şekilde tutulması, sonradan işlenebilir halde saklanması için veri mühendisleri gereklidir.

Makine öğrenmesi mühendisi: Yapay zeka konusunda, bir veya birkaç alan üzerinde, çok ileri düzeyde teorik ve uygulama bilgisine sahip, structured ve unstructured veri üzerinde deneyimli, yapay zeka üzerine ilgili makaleleri okuyup formülize edebilen, matematik ve istatistik bilgisi yüksek düzeyde ancak veri modellemesi ve iş bilgisi üzerinde çok bilgili olmayan (bilgili olmalarına gerek bulunmayan) uzmanlardır.

Deep Learning Engineer şeklinde de ifade edilebilir. Üç tane temel alt alanı vardır (görüntü, ses ve metin). Ancak bunun haricinde, makine öğrenmesi alanı çok geniştir: robotikten tutunda, network analitiğine, route finding'den tutunda, Seq2Seq transformasyonlara....

İş zekası uzmanı: Grafiğin gösterdiği üzere, İş zekası uzmanı, iş bilgisini iyi bilen, çoğunlukla, ilgili işi daha önce operasyon olarak gerçekleştirmiş (ilgili işte çalışmış), yapılan işteki anormal durumları bilen, genel olarak sınırlı istatistiksel bilgiye sahip olan uzmanlardır. Yapay zeka tekniklerini ve veri tabanı modellemesi uzmanlık alanları değildir. Yapılan işe etki eden faktörleri, indikatörleri iyi bilmektedirler.

İş zekası uzmanı ülkemizde şu an en çok çalışanın olduğu alandır. Aslında her (veri bilimi mesleklerinde) meslekte bir iş analisti olmak gerekmektedir. Bu nedenle, biraz ara (geçiş) bir alan gibi de görünebilir.

Veri analisti: Veri analistleri, modellenmiş bir veri tabanındaki verileri, iş birimlerindeki iş uzmanlarından aldıkları donelere ve bilgilere göre inceleyen, veri türleri ve analizi konusunda bilgili, istatistik ve matematiksel bilgilendirme araçları ve “açıklayıcı istatistikler” konusunda deneyimli, analitik – yapay zeka yöntemleri konusunda (genel – çok temel) bilgiye sahip ancak bilgisayar bilimleri konusunda bilgileri kısıtlı olan uzmanlardır.

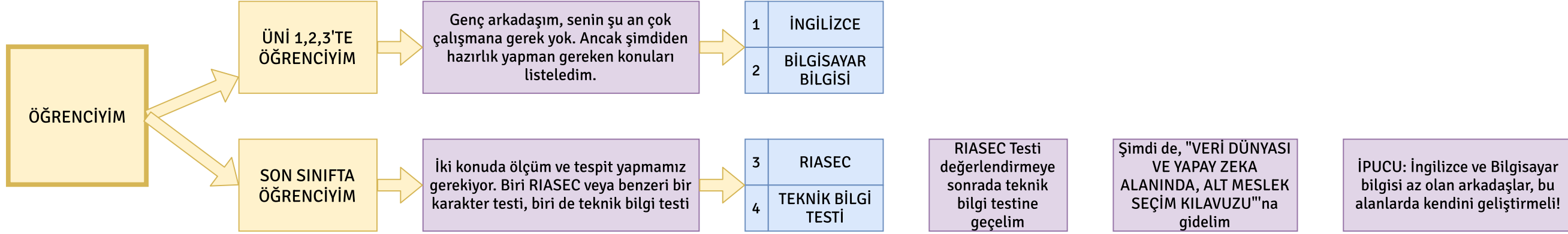
Veri görselleştirme uzmanları da hemen hemen aynı işi yapmaktadır. Sadece biraz daha fazla sunum yeteneğiyle...

Veri Bilimci: Veri Bilimci, bu bahsedilen dört uzmanlık alanının hepsiyle ilgili olan, her biri hakkında bilgili ve deneyimli olan, ancak her bir alandaki uzman kadar detaylı bilgiye sahip olması gerekmeyen uzman kişidir. Daha açık bir ifadeyle, veri bilimci bu dört alanın her birinde eşit miktarda deneyim ve bilgi sahibi olmak zorundadır ancak, her bir alanda, o alanın uzmanı kadar bilmek veya tecrübeli olmak zorunda değildir.

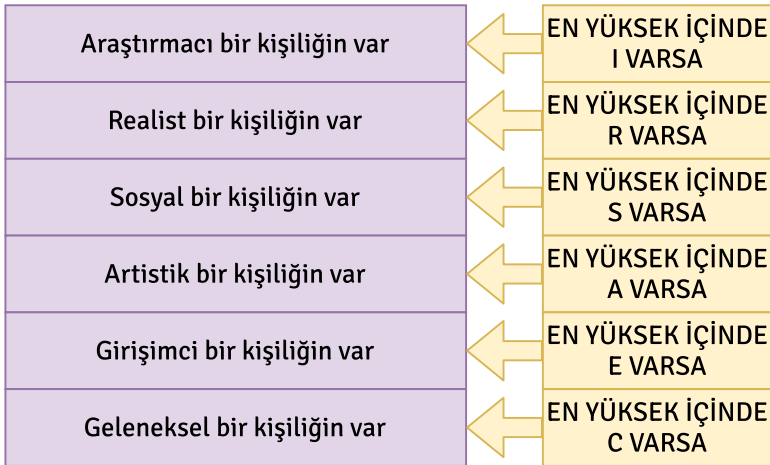
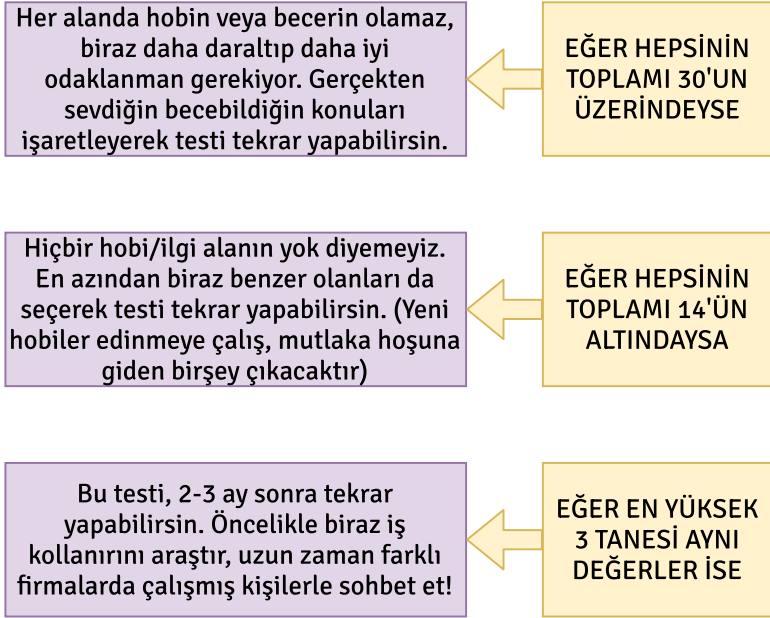
Küçük firmalarda veya firmalardaki analitik bölümlerinin küçük olduğu durumlarda, Veri Bilimci kavramı içinde, yukarıdaki mesleklerin hepsi yapılır.

ÖĞRENCİLER İÇİN VERİ DÜNYASI VE YAPAY ZEKA ALANINDA MESLEK SEÇME HARİTASI

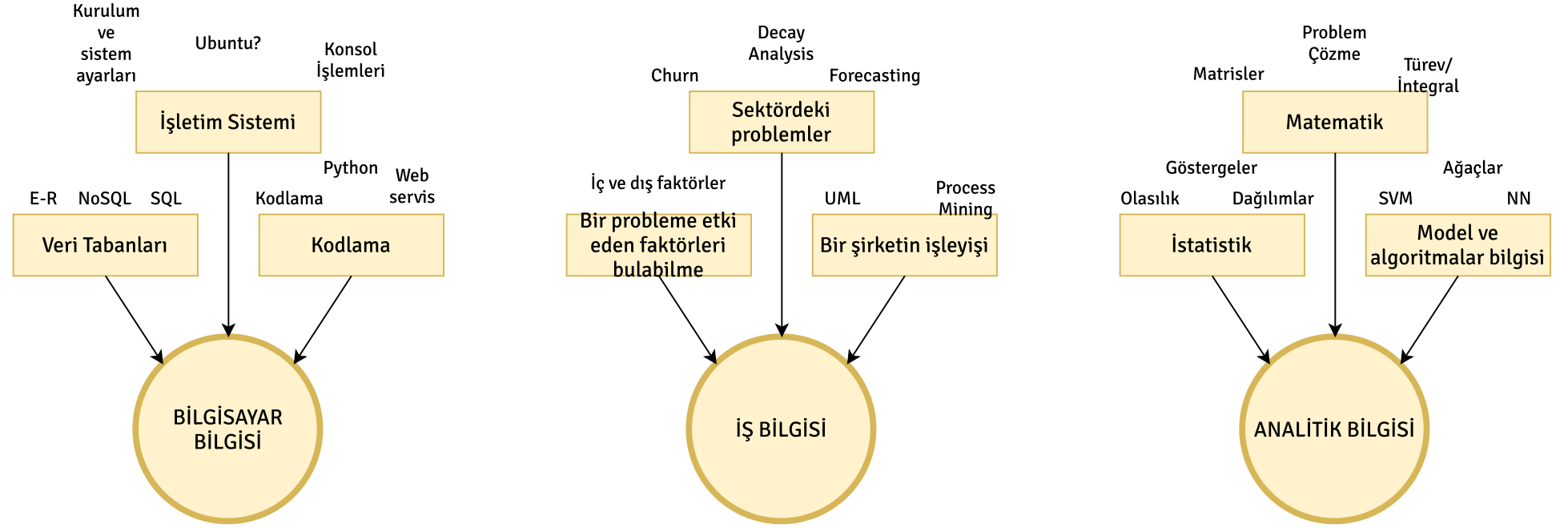
YAPAY ZEKA FAKÜLTESİ - EMRE YAZICI



RIASEC TESTİ DEĞERLENDİRME



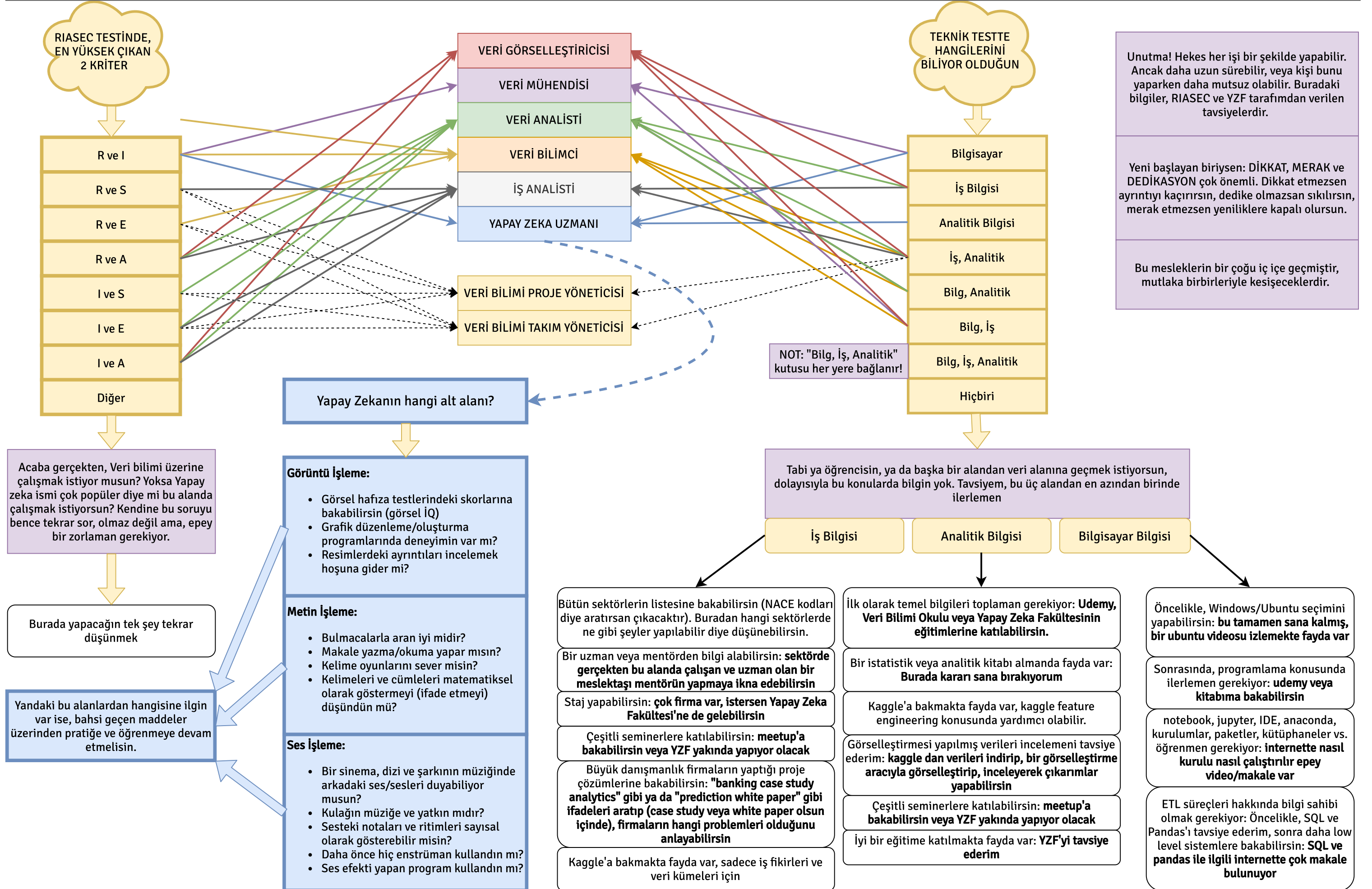
TEKNİK BİLGİ DEĞERLENDİRME

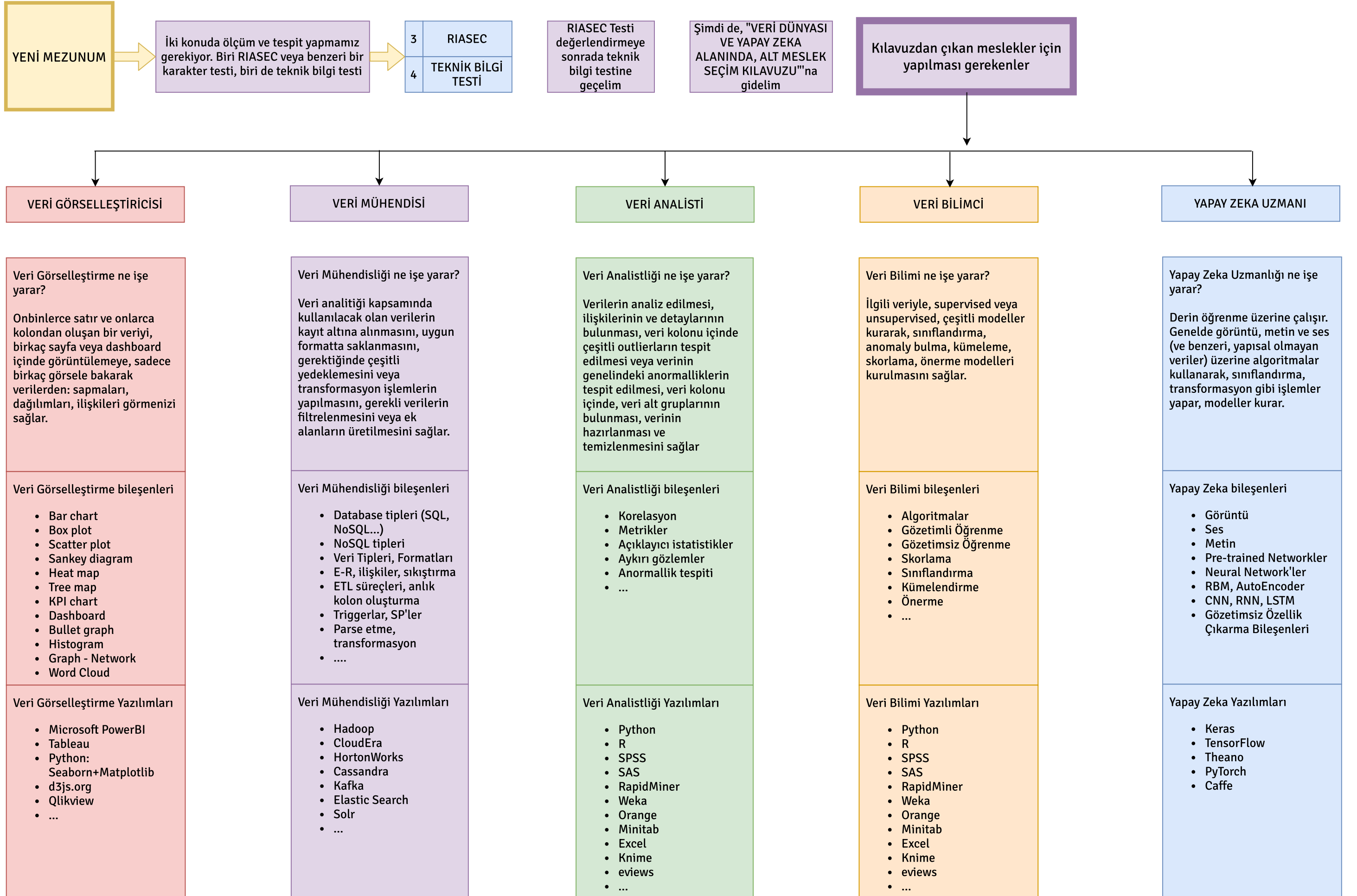


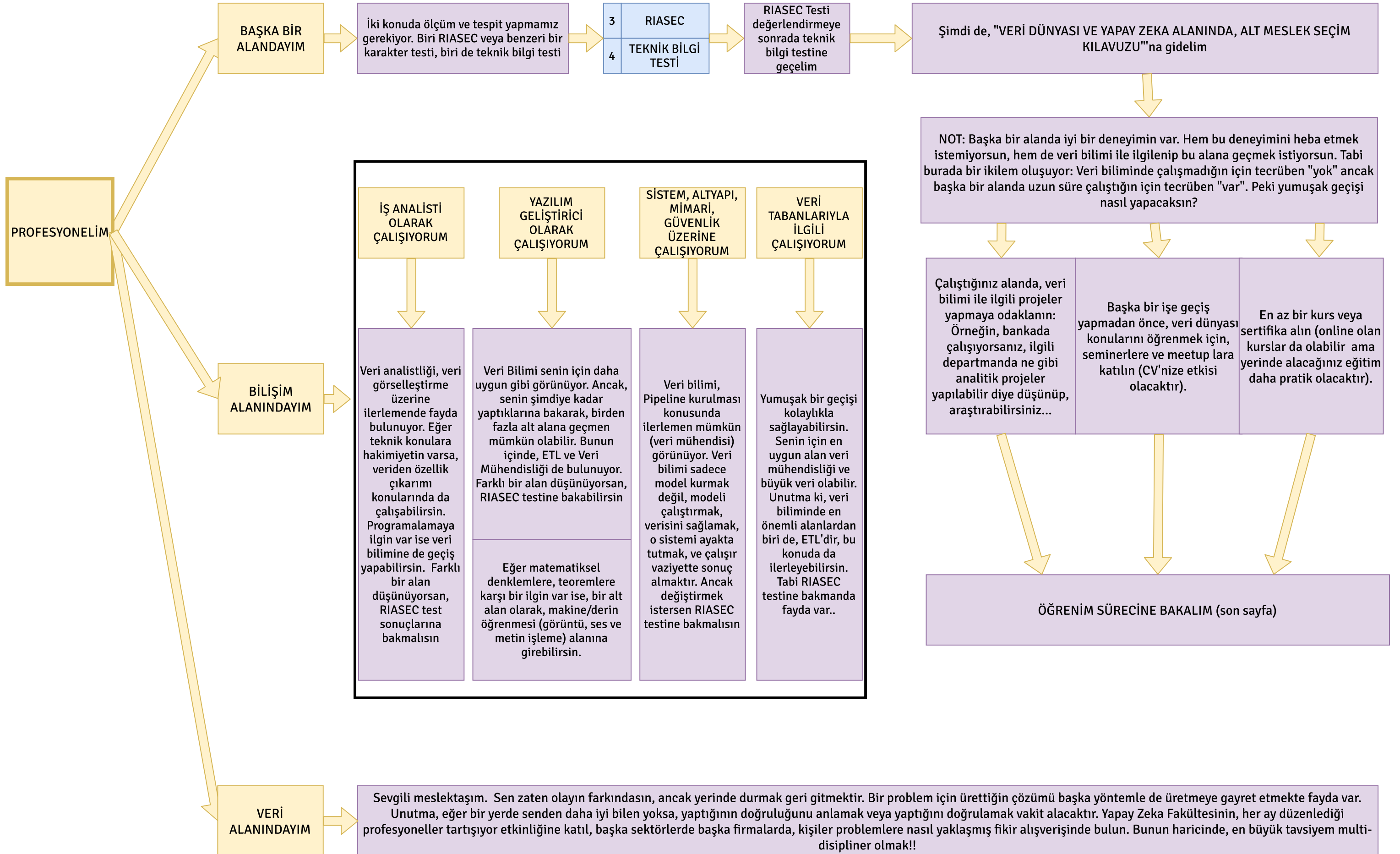
1	İNGİLİZCE	İngilizce, araştırma yapabilmek, yöntem öğrenebilmek için çok önemli. İnternette çok kaynak, çok problem ve çok çözüm bulunuyor. Bunların çoğu da ingilizce dilinde. Bu nedenle, İngilizce çok önemli. Bununla birlikte, google'da arama teknikleri konusundaki ipuçlarına da bakmayı unutmayın.
2	BİLGİSAYAR BİLGİSİ	Yapacağınız bütün çalışmalar, bilgisayar üzerinde olacağı için, öncelikle, işletim sistemleri konusunda iyi bilgiye sahip olmanız gerekiyor. Genel terimlere de vakıf olmanız gerekiyor. Bunun için, Wikipedi bile güzel bir kaynak olabilir. Not: Şahsi ve profesyonel fikrim, Windows'tan kurtulun, siz geliştirici uzmansanız, "Unix" tabanlı bir işletim sistemi kullanın, herşey çok daha kolay. Bunun haricinde, temel bilgisayar bilgisi konusunda (Spreadsheet, Word processor ...) kendinizi mutlaka geliştirin.
3	RIASEC TESTİ	Riasec testi, hangi tip işlerde ve hangi tip pozisyonlarda çalışırsanız daha mutlu olursunuzu gösteren bir testtir. Bu testin, "iyisi/kötüsü" bulunmamaktadır. (https://www.hawaiipublicschools.org/DOE%20Forms/CTE/RIASEC.pdf) adresinde bulunmaktadır. Google'a "riasec test pdf" yazınca "Holand Code" diye de çıkabilir.
4	TEKNİK BİLGİ TESTİ	Yukarıdaki test aslında teknik ve iş konusunda ne kadar bilgili olduğunuzu belirttiğin bir anketir.

İngilizce nasıl öğrenirim:

İngilizce düşünmeye başlayabilirsin. İçinden birşey düşünürken ben "bu"nu nasıl söyledim diye pratik yapın (speaking). Eğer listening i geliştirmek istiyorsan; ingilizce filmleri veya dizileri ingilizce alt yazı ile okumamı tavsiye ederim; böylece duyduğunun doğru olup olmadığını kulak çok çabuk öğrenir. Writing geliştirmek için türkçe - ingilizce kelimeleri yazıp ezberlemeye çalışıyoruz ya; onu çift yönlü yapmakta fayda var. hem ing tur yönünde hem de tur eng yönünde Reading in kolay bir yöntemi; öncelikle tanımını veya içeriğini bildiğin bir konuyu okuman, böylece yeni bilinmeyen kelimelere veya ifadelere takılmıyorsun. wikipediden tanım okumak çok faydalı olur







EĞER VERİ TABANI UZMANI İSENİZ

- Verilere sadece “bytelar” olarak (veya diskte kapladığı yer olarak) değil, manasal olarak bakın. Veriler arası ilişki, veri tabanlarında kullandığınız “relationship”lerden farklıdır.
- Veri tabanında “raw” olarak duran veriler yerine Runtime’da ürettiğiniz değişkenler (Derive değişkenler, “getdate() – dateofbirth”, “Amount/USDRate”) üzerine ve “kolon üreten” sub query’lerle daha çok pratik yapın. SELECT a, b, c, d, (SELECT STDEV(Amount) FROM X WHERE X.a = a) as e FROM Y
- Bir tablonun verisinin tamamını almadan sadece özetitle, o tabloyu tekrardan oluşturmayı deneyin!! (bu örneği gerçekten yapmayı deneyin).
- NoSQL’e bir bakış atın.
- Analitik çalışmalarda, sadece satırlara değil satırların içindeki hücrelere de ulaşp değiştirmeniz veya iteration yapmanız gerekebilir. Cursor’lardan daha iyi yapıları öğrenmek gerekmektedir.

EĞER VERİ ANALİSTİ İSENİZ

- Veri bilimci olmak için, doğru yoldasınız, ancak olaylara biraz daha bilimsel bakmalısınız. Bunun için, belirtilen konuları okumanız faydalı olacaktır: İstatistik, Olasılık, Exploratory data analysis
- Bir analizi veya makine öğrenmesi örneğini bir siteden kopyalayp yapıştırarak çalıştırmak (büyük firmaların sitesinde olsa dahi), sizi veri bilimci yapmayacaktır, o yüzden, “ben biliyorum, bunu yapmıştım” yerine, “ben biliyorum, bunu anlamıştım, kendim implement etmişim” demek amacınız olmalıdır.
- Sadece İş gereksinimlerine odaklanmayın, ki, “amacımız sadece bu problemi çözmek, metodu veya yöntemi anlamak değil” şeklinde yaklaşımları benimsemeyin. Unutmayın, bir iş gereksinimini çözmek önemlidir ve o an için, metodun içeriğini veya nasıl çalıştığını bilmek önemli görünmeyebilir ancak metodların veya yaklaşımların çalışma yapısını anlamak; problemin tipini, başarı oranınızı, zamanınızı değiştirebilir.
- Eğer bir projeyi – çözümü bir yerden alıp kendi probleminizde direkt olarak uygulayabiliyorsanız, işte “o” çözüm doğru çözüm değildir!

EĞER VERİ BİLİMCİ İSENİZ

- Bildiklerinize çok güvenmeyin, her gün yeni metotlar, yeni yaklaşımlar yeni çözümler ve hatta yeni problemler çıkmakta...
- Araştırma ve yeni yaklaşımlar öğrenme, “White paper” okuma, birinci amacınız olmalı.
- Bu kitaptaki her bir adımı daha derin (sprial bir şekilde) öğrenmeniz ve farklı konular üzerinde problemleri farklı yöntemlerle çözmeniz gerekmektedir. Yani bir projeyi “ben bu projeyi anlayarak” yaptım demek ile, “bu projeyi anlayarak yaptım ancak bu 3-4 yöntemle veya yaklaşımla daha yapılabilir” arasında fark vardır.

EĞER YAZILIM GELİŞTİRME UZMANI İSENİZ

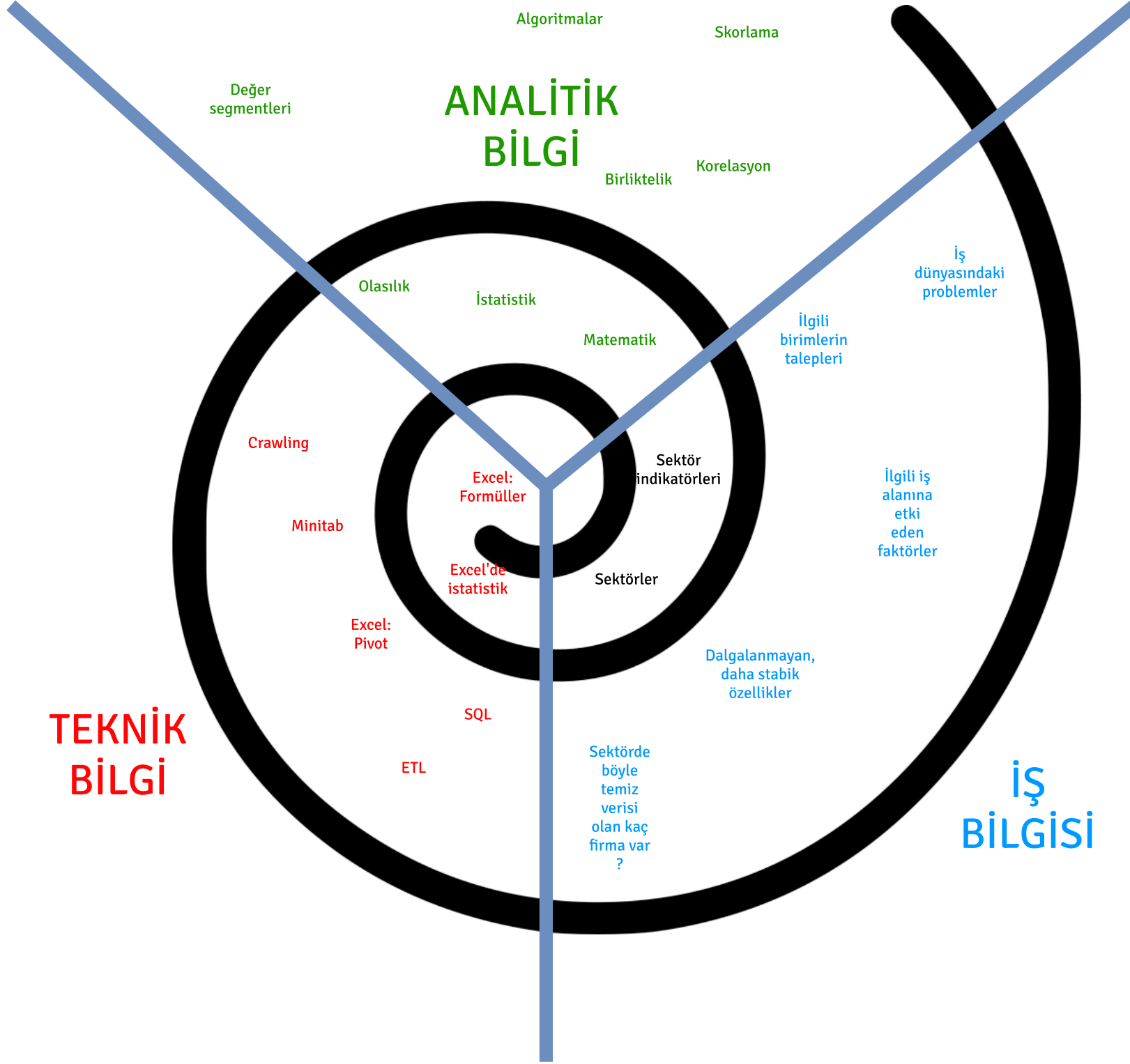
- Veri tabanı bilgilerinizi geliştirin, veri her şeydir.
- Olaylara sadece 1 ve 0 olarak bakmayın, “her şeyin bir kuralı vardır”, “bir fonksiyon yazayım ... hesaplayıp döndürsün”, “if then else”, gibi bilgilerinizi unutun. Analitik dünyada, her şey “muğlaktır” – “siyah beyaz değil de gridir”. Çünkü yazılım geliştiriciler, projeleri tamamen doğru – stabil çalışmak ve verilen iş kurallarına göre sonuç elde etmek için tasarlarlar. Analitik ise, veriden beslendiği için bazen çok garip sonuçlar verebilir. Daha esnek olun, her zaman doğru veya kesin sonuç çıkmayacaktır veya kurallar her zaman sert – kesin olmak zorunda değildir. Saçma ve garip durumlara alışık olun.
- Matematiğiniz zaten iyidir (programlamadan ötürü) ancak istatistik bilgilerinizi geliştirin (tanımlar bölümünü iyice irdeleyin).

EĞER İŞ ANALİSTİ İSENİZ

- Bu pozisyonun pozitif yanı analiz ve problem inceleme konusundaki becerinizin olması, negatif yanı ise (bazıları için) teknik-istatistik ve istatistikten biraz uzak olmanızdır. Analiz ile analitik farklı terimlerdir.
- İleride göreceğiniz “Adımlar” içinde “problem, tipleri analizleri, limitler, çıktılar, v.s.” gibi konuları uygulanmak ve öğrenmek sizin için kolay olacaktır.
- “Veri” leri analiz ederken, çok derine girmemeniz gerekmektedir (çok fazla veri olduğundan, ayrıntılarına girerek kaybolabilirsiniz).
- Karşılaştığınız iş problemini iyi tanımlıyor olabilirsiniz ancak, bazı sorunların çözümleri kesin olmayabilir ve ayrıca problemi çözerken, bazı teknik kısıtların olduğunu da hatırlamak gerekmektedir (örneğin, “bir x sorununu”, “şu” şekilde çözebiliriz diyebilirsiniz ancak o metotun real-time çalışması 5 saniye sürebilir ve sizin anlık cevap vermeniz gerekebilir. Bununla birlikte, verilerinizin doğru olmayabileceğini – yeterli olmayabileceğini ... de hesaba katmanız gerekecektir.).
- Optimizasyon konusuna bakmanızda fayda bulunmaktadır.
- SQL becerinizi ilerletmeniz süreç için size fayda sağlayacaktır.

EĞER BUNLARDAN BİRİ DEĞİLSENİZ

- İstekli olun, kavramları veya sistemleri gözünüzde büyütmeyin, ama küçümsemeyin de. Olaylara KISS ve SIMPLE ile yaklaşın
- Kaggle’ın forumlarına göz atmanız size fayda sağlayacaktır (özellikle başlıklarda “EDA” kelimesi geçenlere).
- Temel manada (uygulamadan bağımsız olarak) aşağıdaki konulara göz atın.
 - Programlama
 - Veri tabanı
 - İstatistik



Veri dünyasındaki konulara hakimiyet için öğrenim süreci bir spiral şeklinde gelişmelidir. Çünkü, öğreneceğiniz konseptler birbirleriyle mutlaka (az da olsa) ilişkilidir. Bir tanesini öğrenmeden diğerine geçince, bir projeyi sadece "o bakış" açısından değerlendirebiliyor olursunuz ki, bu sizin daha geniş ve detaylı düşünebilmenizi engeller. Bu nedenle, spiral öğrenme ile ilerlemek mantıklıdır. Peki nedir bu spiral öğrenme?

İlk olarak, en güvendiğiniz veya size en kolay gelecek olan (3 alandan biri: iş bilgisi, analitik bilgi, teknik bilgi) birini seçin, ve o alanla ilgili küçük bir bilgi öğrenin.

Örneğin, teknik bilgiyi seçtiniz, yani bilgisayar bilgisi de diyebiliriz. Basit bir konuyla başlayalım: Excel (Spreadsheet). Excel'i çok iyi bilmiyorsanız, formülleri, hesaplamaları, vs öğrenebilirsiniz (çok basit konular!)

Şimdi ikinci adıma geçelim, sonraki adımda da diyelim ki, İş bilgisi alanından devam edeceğiz. Şimdi daha önceden öğrendiğiniz, formüller ve hesaplamalarla, iş bilgisini pekiştirerek öğrenin. Örneğin, bir sektörün, (veya satış bilgisinin, müşteri bilgisinin) verilerini alıp, bu Excel'den bildiğiniz, formülleri ve hesaplamaları da kullanarak çeşitli indicatörler oluşturun (veya internetten okuyun). Yine örnek olarak, sektörlerin karlılığı, sektördeki firma sayısı, sektördeki geçen seneye göre olan değişimin olduğu veriyi Excel'de inceleyerek, geçen senedeki, sektörün karının sektördeki firmalara bölümüyle bu seneki oran arasındaki farka bakın. Yine basit bir indicatör olsa bile, hem iş dünyasından bir indikatör öğrenmiş oldunuz hem de öğrendiğinizi uygulamış olduğunuz.

Şimdi sonraki adım olarna analitik bilgiye geçelim. Burada kendimizi biraz daha açıyoruz, ve standard sapmayı öğreniyoruz. Sonra yukarıda bahsettiğimiz veride, sektörlerin her bir yıl içindeki bir verisinin standard sapmalarına bakıyoruz ve bazı sektörlerdeki, standard sapmanın çok büyük olduğunu görüyoruz, bu da, o sektörün verisinin çok değişken olduğunu ve bu değişken ile çalışırken başka çeşit normalizasyonlar yapmamız gerektiğine işaret ediyor.

Sonra yine başa - bilgisayar bilgisine dönüyoruz, çeşitli dış faktörleri indirip başka excel dosyalarıyla birleştiriyoruz (örneğin enflasyon verisi) ve her birini kendi içinde yüzdelik dağılıma göre gösteriyoruz.

.....

Bu şekilde, her bir adımda, biraz daha fazla, geniş, derin bilgi öğrenebiliriz. Ayrıca, bir önce öğrendiğimizi de pratik ettiğimiz için öğrendiklerimizi unutmamız engellenmiş olur.

Bu döngüyü birkaç kere yaptıktan sonra, artık ilerlemek istediğiniz noktaya doğru gidebilirsiniz. Birkaç kereden sonra, artık genel olarak her alanda bilginiz bulunacaktır ve bundan sonra diğer alanlarda çok detaylı bilgiye her zaman ihtiyacınız olmayacaktır.

Yapay Zeka Fakültesi olarak, bu süreçte seni yalnız bırakmıyoruz.

Bize katıl, ipuçlarından, projelerden yararlan.