

BİLGİSAYAR MÜHENDİSLERİ KURULTAYI
10-11 ŞUBAT 2018, BOĞAZIÇI ÜNİVERSİTESİ-İSTANBUL
Yapay Zekâ ve Bilgisayar Mühendisinin Sorumluluğu

Prof. Dr. Cem SAY
Boğaziçi Üniversitesi Bilgisayar Mühendisliği Bölümü Öğretim Üyesi

Merhabalar. Çok teşekkür ediyorum. Bu toplantı çok harika oldu. Özellikle bu salona alındığımı duyunca daha da çok sevindim. Neden? Hem tarihi bir salonumuz hem de yapay zekâ açısından şöyle bir enteresanlığı var. Bilmem fark ettiniz mi, ama ben kızımınla gittiğimde ikimiz de fark ettik. Türkiye'de yapay zekâyla ilgili çok fazla film yok; en önemlisi bu (*film: Arif V 216*). Dikkat ettiyseniz bu filmin bir sahnesi de bu salonda çekilmiş [Yansı-1]. Bir profesör var; insanlar iyi midir, kötü müdür konusunda araştırma yapıyor. Onun ofisi de -Boğaziçi'ndeki bütün hocaların ofisleri böyle olsa keşke- bu salon. Bu salon kullanılmış o iş için. O yüzden çok uygun olmuş sağ olun.



Yansı-1: Boğaziçi Üniversitesi Kütüphanesinde Çekilen Yapay Zekâyla İlgili Film

Yapay zekâ, bu konuşmanın tanıtımında da yazıldığı gibi günümüzün gözdesi. Hoş, aslında bu resim günümüzden değil, 20 seneden fazla oldu; ama IBM'ci arkadaşlarımız da var, o yüzden onlara da göstermeden edemiyoruz. Gerçekten de başarıdan başarıya koşuyor. Dünya satranç şampiyonunu yenişi hâlâ hatırlarda... Daha yeni resimlerimiz de var. Bu, geçen seneden. Dünya “go” şampiyonunun Google'ın yapay zekâsı karşısında girdiği hal [Yansı-2]. 19 yaşında dünya “go” şampiyonu, hiç yenilgisi yok ama Google'ın DeepMind şirketinin derin öğrenme yoluyla kendi kendine bu oyunu öğrenen makinesinin karşısında 3-0 yenildi Çinli şampiyon.



Yansı-2: Dünya “go” Şampiyonu, Google'ın Yapay Zekâsı Karşısında

Bu kadıncağızı Youtube'da belki görmüşsünüzdür. Torunu buna TESLA araba satın almış. Araba sürmeyi bilen bir hanımefendi ama hiç kendi kendine giden arabanın direksiyonunda bulunmamış tabii [Yansı-3]. Youtube'da “autopilot grandmother” diye aratırsanız çok güzel, heyecanlı deneyimini siz de paylaşabilirsiniz.



Yansı-3: “Autopilot Grandmother”

Bunlar bizim, yani benim gençliğimde bilgisayar mühendisliği eğitimi sırasında hayal gibi görünen şeyler. Bunların hepsi birer birer oluyor. Başka şeyler de var. Böyle birçok örnek var ama bebekli olduğundan bu haberi seçtim. Bu çocukcağızın nefes yolları enfeksiyonu var. Nefes alamaz hale geliyor. Annesi koşuyor, telefonu bırakıp çocuğa müdahale ederken bir yandan da telefona bağırarak ambulans çağırmasını söylüyor Siri'ye. Siri de ambulans çağırıyor. Annesi de Apple'a teşekkür mektubu yazıyor. Bunu sağlayan ses tanıma teknolojisi de benim öğrenciliğimi bırakın işte bayağı yakın bir zaman önce pek mümkün olduğunu göremediğimiz teknolojilerdendi. Yine bunu da sağlamış durumdayız.

Facebook'un intihar edecek gibi olan kullanıcılarını anlayıp ondan sonra da duruma müdahale etme servisi var. Çünkü maalesef böyle olaylar Facebook'ta maalesef sık sık oluyor. Özellikle de bu canlı yayın olanağı olduğu için. Hem başka kullanıcılardan gelen uyarıları hızlı bir şekilde işleyip hem de yapay zekâmsı teknikler kullanarak yayınları otomatik olarak analiz edip böyle bir durum varsa otomatik olarak birilerini uyarıp ya kişiyle doğrudan temasa geçip ya da onun arkadaşlarını -çünkü biliyorsunuz arkadaşlarınızın kim olduğunu Facebook sizden daha iyi biliyor- ya da en yakındaki polisi uyarıp müdahale etme sistemini kurmuş durumda Facebook. Avrupa sınırları haricinde. Çünkü Avrupa kanunları, burada kullandığı bilgileri kullanmasına izin vermediğinden intihar etmek üzere olan Avrupalılara bu yardımı yapmıyor; ama dünyanın geri kalanında olanlara yapıyor.

Sonra bu var, IBM'in Watson'u... Biliyorsunuz, haklı olarak övünüyorlar; çok geniş tıp literatürüne hâkim olabiliyor. Okuyup insanların gözünden kaçabilecek birtakım bağlantıları görebiliyor. Bu da övündükleri olaylardan biri; Watson'un bilgi yarışması kazanmasından kalan bir resmi. Bu haberde anlatılan olay şu: Japonya'dan bir hanım var, kanser hastası. 7-8 senedir lösemi tedavisi görüyor, ama bir türlü iyileşemiyor. Bir de Watson'a sorulmuş diyorlar. Watson, 10 dakika içinde bütün tıp literatürüyle kadının genetik analizlerini karşılaştırarak gözden kaçan bir şeyi buluyor. Meğerse kadında bir değil, iki tane değişik lösemi varmış. Watson, “Bakın bu da varmış,” diyor ve onun tedavisine de başlıyorlar. Kadının durumu iyileşiyor.

Bu nedenlerle gerçekten de yapay zekâ, kamuoyunda son zamanlarda iyice bir rüzgâr estirmeye başladı. İnsanlar işleyişini çok iyi bilmedikleri teknolojiyle değişik bir ilişki içine giriyorlar. Ya fazla değer veriyorlar “Artık bunu da yaptığına göre her şeyi yapar” diye düşünüyorlar ya da biraz korkmaya başlıyorlar. “Eyvah. Bunu da yaptığına göre her şeyi yapar” türünden bir korku oluyor...

Özetle -zaten Batı toplumunun kültürel altyapısı buna uygun- kaygı duymamıza yol açan durumlar da var. Elon Musk -özellikle Nick Bostrom diye bir felsefeci var; onun yazdıklarından etkilenerek- bildiğiniz gibi bu işin sonunun iyi olmayabileceğine dair şeyler söylüyor. Bir yandan da kendisinin yapay zekâ şirketi var, nasıl oluyorsa... Stephen Hawking, biliyorsunuz artık vaktini bu gibi konularda felaket tellallığı yaparak geçiriyor. Benim bir arkadaşım var, “Sen bu işlerle uğraşma, bunun üstüne su dökülürse hepimizi keser,” falan diye söylüyor. Böyle bir kaygı boyutu da var işin. O yüzden bilgisayar mühendisleri olarak bu işin nasıl çalıştığını, ne olup ne bittiğini anlayan, duruma hâkim kişiler olarak bize ne sorumluluklar düşüyor?

Yanlış anlamaları düzeltereğiz tabii, mümkün olduğu kadar... Çünkü yanlış anlamaların devam etmesi gerçekten bu teknolojinin anlamlı kullanımlarının da değersizleşmesi, onların da kamuoyu ya da karar vericiler tarafından yanlış değerlendirme sonucu önünün kesilmesine neden olabilir. Ürünümüz hakkında aşırı reklama kaçmamamız için ürünümüzün tam olarak şu andaki yapay zekâ teknolojisinin eksikliklerine, hatalarına hâkim olmamız, onu anlamamız ve mümkünse onu iyileştirip hakikaten iyi bir hale getirmemiz lazım. Konuşmamda, görebildiğim kadarıyla şimdi yapay zekâ teknolojisinin eksikleri, sıkıntıları nelerdir, onlardan söz etmeye çalışacağım.

Kötüye kullanılması olasılığı olan bir teknoloji. Aslında her teknolojide bu var... Ben, yapay zekâyla ilgilendiğim için onun çeşitli kötüye kullanım olasılıklarından söz edeceğim. İleride, kötü niyetli olmasa da bu şekilde gidersek, toplumumuzu bu hale getirirsek o durumda başımıza neler gelebilir konusunda öngörülerde bulunmaya çalışacağım.

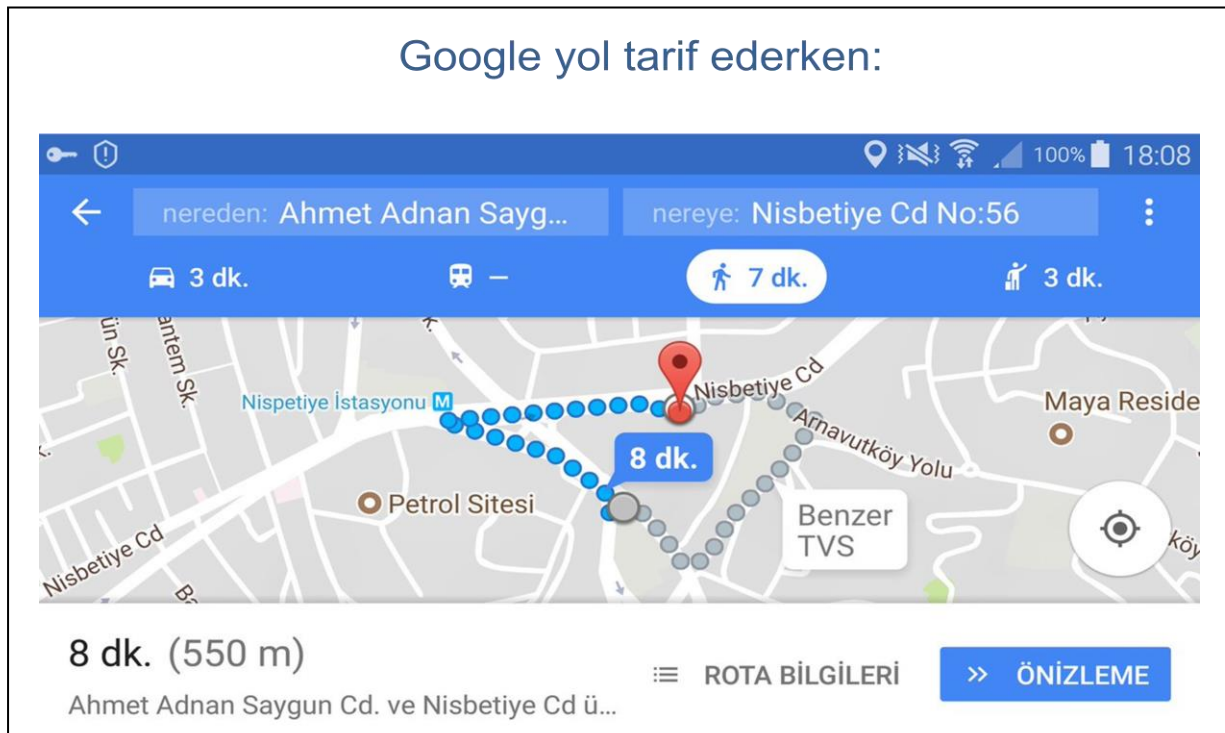
Öngöründe bulunmak aslında çok zor bir şey. Mümkünse bilgisayar mühendisleri olarak yapay zekâyı da becerebilirsek dünyayı da şimdi olduğundan çok çok daha iyi bir hale getirmek; esas hedefimiz bu herhalde. Biz bu işi neden yapıyoruz? Mümkünse bunu yapalım, tarzında bir şeyler söyleyeceğim; ama belki bunların hepsinden daha önceliği olan şey, bilgisayar mühendislerinin “ne oluyor” diye bu konuda sürekli kendilerini güncel tutmaları gerekiyor. Bu çok önemli. Ben yetişemiyorum yani... Bunun bir sistemini kurmamız lazım. Takdir edersiniz bu meslek öyle bir meslek. 1987'de mezun oldum. O bilgiyle dursaydım şimdi “Bilgisayarlar yakında satrançta insanları yenebilir” filan noktasında kalmam gerekecekti. Sürekli kendimizi güncellememiz gerekiyor. Roma Hukuku olsaydı keşke konumuz. Hiçbir yenilik olmadan emekliliğe kadar idare edebilirdik...

Birkaç yanlış anlamadan bahsedeceğim. Ta benim doğduğum yıl -şimdi hangi yıl olduğunu söylemeyeyim- şöyle bir problem çıkıyor... Eliza diye meşhur bir program. Şimdiki “chatbot”ların büyük büyük annesi. İlk “chatbot”. Çok çok basit bir şekilde çalışıyor. Bir cümleyi alıyor, sadece “ben”leri “sen” yapıp geri veriyor. Bazı psikologlar öyle çalışıyor ya... “Ben bugün çok mutsuzum” diyor hasta, “Sen bugün çok mu mutsuzsun?” diye yanıt veriyor; yani tamamen o. Çok basit bir değişiklik. Türkçede yapmak biraz daha zor ama oyun olsun diye ben de yıllar önce Türkçesini yapmıştım. Neyse bu konuşmayı izlerseniz, zaten foyasının ne olduğu çok kısa zamanda meydana çıkıyor. Bu bile -Joseph Weizenbaum adlı bir bilim adamı yapmış bunu- hakikaten ciddi ciddi bir bilimsel dergide yayımlanmış. O zaman için iyi bir iş. Adam artık endişelenmeye başlıyor, çünkü gerçekten bunu konuşturduğu, Eliza ile temasa geçmesini söylediği insanlar “Sen çıkar mısın? Ben özel bir şey konuşacağım” filan diye gerçekten onun (*Eliza'nın*) anladığı varsayımını hoyratça

Bu, Alice için mesela... Alice'in üç kitabı var, kitap başına 1 dolar değer biçiyor kendi kafasında. Öte yandan Bob'un üç kitabı var, ama o kitaplara hiç değer vermiyor. İkisinin de değişik değer fonksiyonları var. Birbirleriyle takas edecekler. Bu takası “ben sana şu kadar vereyim, sen bana bu kadar ver” diye pazarlık şeklinde yapacaklar. Sonunda ikisi de elde ettikleri fonksiyona ulaşmaya çalışacaklar. Böylelikle pazarlık yapma yeteneklerine sahip olan “bot”ları binlerce defa kendi aralarında pazarlık yapıp daha becerikli olanı bir sonraki jenerasyona aktararak üretmeye, bir yandan evrilmeye çalışıyorlar. Fakat işte o işi yaparken “Doğru dürüst İngilizce konuşmaya devam edin” anlamına gelen değeri oraya eklemeyi unuttunca böyle oluyor. Kendi aralarında bu işe uygun, birbirlerinin anlayabildiği ama insanların anlayamadığı bir dile doğru evriliyor. Bu, dilin nasıl evrildiğine dair ilginç bir bilgi. Dilbilimciler açısından ilginç bir araştırma alanı olabilir; ama nasılsa saçma bir şekilde medyada yansıtılıyor. Bunları düzeltmeye çalışmamız lazım.

Yapay zekânın görebildiğim kadarıyla şu andaki teknolojisindeki sıkıntıları, eksiklikleri paylaşmak isterim. Birbirimizle bu konuları yani özellikle “Vay sen benim programımda hata buldun” diye kızmak yerine -böyle yapmıyorlar zaten- özellikle bu mevcut parlak teknolojilerin olabiliyorsa zayıf noktalarını bulup onların üzerine yürümemiz lazım. Bu konuda ciddi akademik çalışmalar oluyor. Oradan birkaç örnek hazırladım.

Benim Google'a eziyet etmekle ilgili bir hobim var. Bu, yürüme yol tarifi. Bu, başlangıç ve bitiş noktaları. Akmerkez denen -buraya sadece 1-2 km uzakta olan- üçgen şeklindeki binanın iki değişik giriş kapısından birinden diğerine nasıl gidilir diye sordum. Tabii doğru yanıt “Kapının birinden gir, içeriden diğerinden çık” olacak. Neyse onu yapamıyor ama şubat itibariyle bunu sorduğumda beni alıp bayağı uzak bir yere götürüp oradan döndürüp -zaten böyle bir yol tarifinde döngü oldu mu saçmalık olduğunu görmemiz gerekiyor, daha kısaltılabilir çünkü- tekrar getiriyordu; ama sonra düzelttiler. Bu düzelmiş hali. Şimdi de nedense önce bir metro istasyonuna kadar yürütüyor gereksiz şekilde. Bu yol taş çatlasa iki dakika sürer; ama en azından şubattakine oranla daha iyi. Elbette ki bunun neden olduğunu anlayabiliyoruz hepimiz. Bu, aslında algoritmanın kabahati değil. Bu, haritanın kabahati. [Yansı-6]



Yansı-6: Yol Tarifi

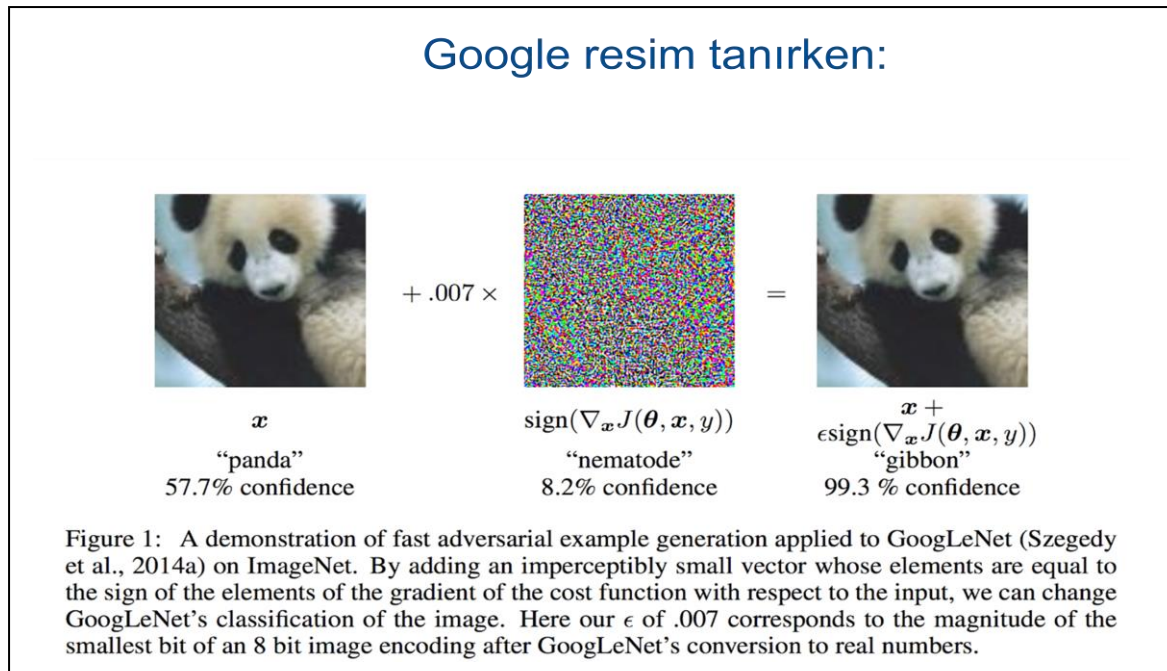
Ama bir de algoritmanın gerçek hayatta tam eklememesiyle, gerçek zamanlı bilgiler alıp almamasıyla ilgili şöyle sıkıntılar olabiliyor. Bu, iki ay öncesinden bir haber. 7 Aralık 2017'de. California'da yangınlar oluyor mevsimsel olarak. İşte onlardan birinde, yine bu Google gibi yol tarif uygulamaları insanları “Bak bu yol boş, buradan git” diye yangın olan yerlere yönlendiriyormuş. Tamam, o da onun kabahati değil. Biz burada mühendisler olarak hep programı savunma çabasındayız; ama gördüğümüz gibi burada ciddi bir sıkıntı var. O bilgilerin gerçek zamanlı entegre edilmesi lazım; eğer siz “Ben gerçek zamanlı yol tarif programı yaptım” diye geziyorsanız... Nisan 2017 itibarıyla bunlar düzeltildi; ama çok güzel olduklarından ben bunu bütün sunumlarımda kullanıyorum.

“I am a .. bir şey” tipi bir şey yazdığınız zaman Google'ın çevirisine -bu uzun yıllar böyle devam etti, şimdi düzelttiler- “Ben o şey değilim” diye çeviriyordu Türkçeye. Yani hiç yoktan bir “değilim” ekliyordu. Pozitif, negatif yapıyordu. Artık bunları yapmıyor. Ben bunu Herkese Bilim Teknoloji dergisinde de yazdım. Ben yazdım ve Google'ın patronu o gün Herkese Bilim Teknoloji dergisini arattırıp hemen düzelttirdi.

Bunlar, çalışırken aldığım birkaç ekran görüntüsü. Siz de deneyebilirsiniz telefonunuzdan hemen. Burada tabii sıkıntı var. “He/She” meselesinde zaten Türkçeye İngilizce arasında sıkıntı var ve orada genellikle yanlışlar oluyor. Mesela hemşireleri “She” diye çeviriyor, öyle şeyler var. Ama burada gerçekten veri azlığından -herhalde- kaynaklanan bir sıkıntı var. “Benim elmam kırmızı” gayet güzel çevrilmiş ama “Senin elman ne renk” çok fena olmuş. “Elman” sözcüğü galiba Azericede bir isim falan ve o dengesini bozuyor. Çünkü bu yeni geçtikleri sistemde -benim yıllarca saçlarımı dökme pahasına çalıştığım- cümleyi özne, şu, şu diye ayır, ondan sonra onları tekrar öbür tarafta kur, falan yapmıyor. Genel olarak bu dildeki şu cümle, diğer dildeki şu cümleye benziyor, demek ki ona benzeyen öbür cümle de birazcık daha bozulmuş bu cümleye benzer diye geometrik dönüşüm yapıyor. Çok komik şeyler çıkabiliyor.

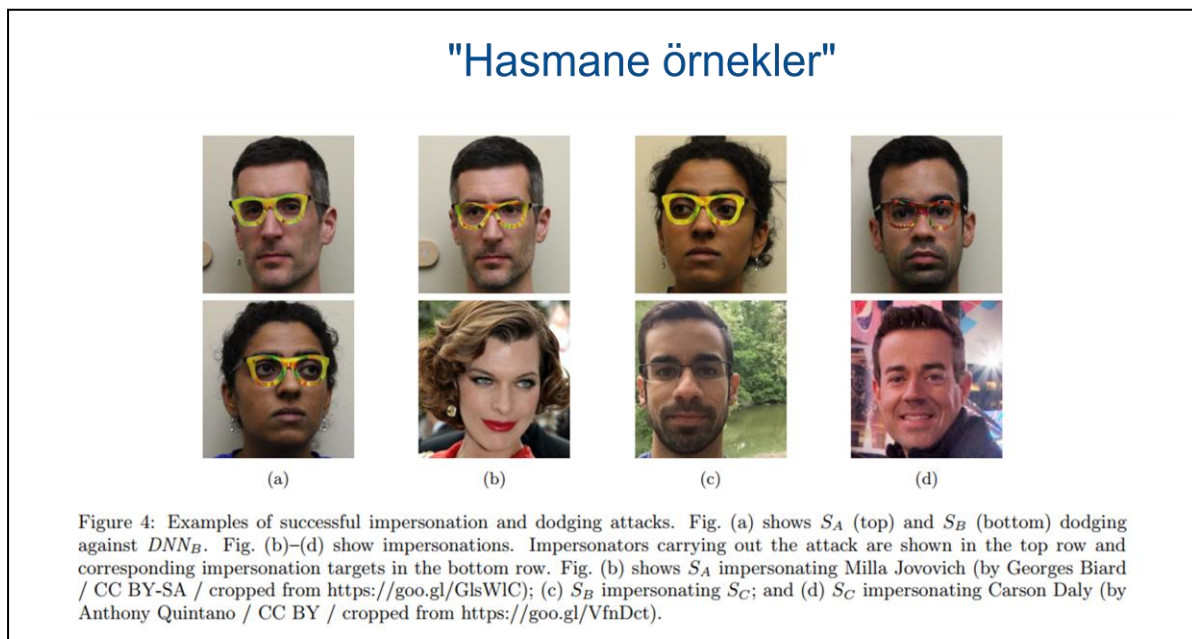
Başka birine bakarsak.. Yani bu gibi işleri.. Yabancılarla gönül ilişkilerine giriyorsanız bence dilini öğrenmeniz lazım. Kesinlikle Google Çeviri üzerinden yapmayın. Şimdi bu mesela, makine öğrenmesinin önemli konularından biri. Çok çok enteresan bir konu. Belki daha önce görmüşsünüzdür. Özellikle sinir ağlarının nasıl çalıştığını biliyoruz. Bunlarla ilgili şöyle bir problem var. Son yıllarda gerçekten örüntü tanımada, demin de dediğim gibi hayal edemeyeceğimiz çok güzel noktalara geldik. Ama eğer sinir ağının nasıl çalıştığını biliyorsanız, onun üzerinde deney yapma şansınız varsa, onun verdiği yanıtı bozmak için -buna hasmane örnekler (“adversary examples”) deniyor- onun yanlış yapabileceği örnekler üretebiliyorsunuz. Nöronlar arasındaki ağırlıklarla çok azıcık oynayarak, oradan bir hesap yaparak.

Solda gerçekten de bir panda resmi var. Ve Google'cıların sanırım birkaç sene öncesinde Imagenet yarışmasını kazanırken kullandıkları sinir ağı. İşte, dal falan da var, ikinci kulağı görünmüyor. O yüzden tam emin olamıyor ama %57,7 güvenle “Bu bir pandadır” diyor. Ama sonra o sinir ağının ağırlıklarında hesapladığımız çok küçük bir, böyle ne olduğu belli olmayan, siyah beyaz noktalardan oluşan bir ağırlık üstüne koyuyoruz onun. Sağdaki resmi, ortadaki resmi koymuyoruz onun binde birini ağırlıklandırarak koyuyoruz. Soldaki iki resmi eklediğimizde çıkan resim yine şu sağdaki. Yani insan gözü için birinciyle üçüncü için hiç fark yok; ama bu sinir ağı açısından bir optik illüzyon yaratmayı başarıyor ve o sinir ağı şu sağdaki resme %99 güvenle bu “Başka bir hayvandır; gibondur,” diyor [Yansı-7]. Bunun bir sürü enteresan örneği var.



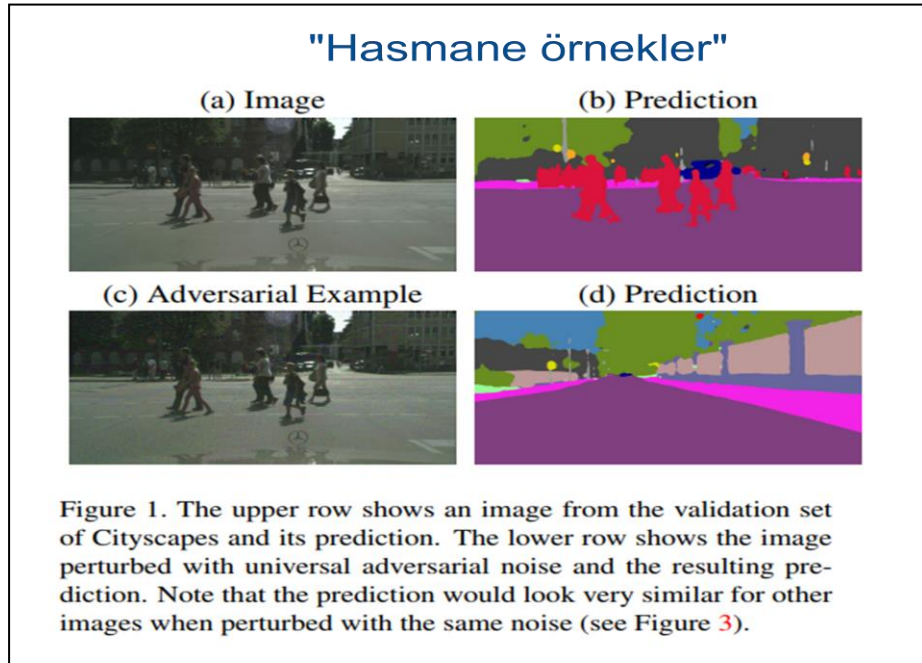
Yansı-7: Resim Tanıma

Bu adamlarınkini çok sevdim. Bu adamlar, yüz tanıma sistemlerini yanıltma konulu makaleler yazan insanlar. Yüz tanıma sistemine kişilerin yüzlerini yüklüyorlar. "Hah, bu Ahmet; terörist, yakalayın," falan... Ya da işte "Giriş hakkı var, kapı açılsın" diye... Bunlar da o sistemleri yanıltıcı desenli gözlükler "print" edip, gözlerine takıp neler olduğunu gösteriyorlar. Soldaki kolondaki iki kişi -bu makalenin yazarları- sisteme tanıtılmış durumdalar; ama o gözlükleri taktıkları durumda Makine onları tanımıyor, başka insan deyip geçmelerine izin veriyor. İkinci kolonda yine makalenin yazarı kendisini Milla Jovovich olarak tanıtmayı başarıyor. Makine yani Milla Jovovich'i de tanımış, onu da tanımış ama üstteki adam o gözlüğü taktığında Milla Jovovich diye tanıyor makine. Yine ikinci kolonda o gözlük sayesinde kadıncağız alttaki genç olarak tanımıyor makine tarafından. Dördüncü kolonda da karşısındaki kimdir bilmiyorum, ama herhalde yine film yıldızı gibi bir şey olması lazım. Sağdaki bayağı Ortadoğulu, koyu tenli arkadaş %16 olasılıkla, ama o az oluyor; Carson Daly olarak geçiyor. Sadece uygun gözlüğü "print" edip takarsanız. [Yansı-8]



Yansı-8: Hasmane Örnekler

Bir de bu, hakikaten kötü uygulama var. Yine birinci kolonda gerçek imaj ve demin anlattığım şekilde onun bizim gözümüzün alamayacağı kadar ama sinir ağını yanıltacak şekilde “perturbe” edilmiş versiyonu var. Sağ kolona bakarsak “Hah burada şu şu yayalar var” şeklinde yayaları görüyor ve araba fren yapıyor herhalde. Ama maalesef ikinci imaj gördüğünüz gibi yayaları silmek için konulmuş. “Ohh, bomboş bir yol var, buradan gazlayabilirsin” bilgisi gidiyor arabaya. Bunun arabayı “hack”leyen kötü kalpli insanlar tarafından nasıl kullanılacağını düşünün. [Yansı-9]



Yansı-9: Hasmane Örnekler

Ve tabii Microsoft'un konuşmaların gözdesi olan örneği... Microsoft, bir Twitter karakteri yapıyor. 19 yaşında genç bir kız. O “tweet”ler atacak, insanlar da yanıt verecek. Böyle böyle derin öğrenme tekniği ile İngilizce “tweet”leşmeyi mükemmel bir şekilde öğrenecek diyorlar 23 Mart 2016'da. Ben sabah bu haberi okudum, “Eyvah” dedim. Ertesi gün, o kadar kötü şeyler söylüyorlar ki çocuğunuzun kontrolsüz şekilde başkalarıyla konuşmasına izin verirseniz o kadar fena şeyler söyleyen bir karakter olmayı öğreniyor ki... Kapatıyorlar. Bir daha da açmadılar o zamandan bu yana.

Bu hataların ne olduğunu anlayıp sıkıntıları düzeltmeye çalışmamız lazım tabii. Bildiğimiz kadarıyla şimdi çalışan teknolojimizin nasıl kötüye kullanılabileceğine dair sorular soralım. Ya da nasıl sıkıntılar yaratabileceğine dair sorular soralım. “Yapay zekâ işsizlik yaratacak mı” meselesi çok konuşulan bir şey. Bazen de hani böyle “Dünyada hiç kimsenin işi kalmayacak, 5 seneye kadar bitti” falan gibi aşırı tahminler de yapılabiliyor. Ben o görüşte değilim; ama zaten ben geleceği göremediğimi daha önce söylemiştim. Önceki teknolojik sıçramalarda uzun vadede işsizlik artmamış. Ekonomistler onu diyor. Bazı işler yok olmuş, ama yerine teknolojiyle ilgili yeni işler gelmiş. Bu sefer de öyle mi olur? Bilemiyoruz. Öyle olsa bile geçiş dönemi diye bir şey var. Ciddi şekilde kendini süren arabaların gelmekte olduğunu ve de gelirlerse de iyi olacağını, yani trafik kazalarının azalacağını, trafiğin çok çok daha iyi bir şekilde, hızlı bir şekilde akacağını görüyoruz. Ama o kadar bin tane veya o kadar milyon tane şoför ne olacak? Üstelik bu şoförler de tam teröristliğe geçiş yapabilecek insan profilinde... Şöyle demek istiyorum yani bekar, erkek. Tam, “Biz Irak'ı işgal ettik, Irak ordusunu komple terhis ettik. Ne kötü gidebilir ki” filan diye

Amerikalıların yaptığını düşünün. Sonra onlar İŞİD'çi olabilir mesela. Bir anda bir grup insanı işsiz bırakırsanız ne olabileceğini düşünmeniz lazım.

Buna karşılık önerilen “evrensel temel gelir” fikrini duymuşsunuzdur. Devlet herkese belli sabit bir şey verecek, yaratılan zenginlik mecburen bunları dengeli tutacak şekilde dağıtılacak. Bu sistem çalışır, eğer herkesin bir oyu olursa tabii. Herkesin bir oyu olmazsa o zaman gerçekten de buna gerek yok. “Bir grup insana gerek yok” noktasına gelmemiz iyi olmaz. Bu gibi durumlarda genellikle ABD'deki at nüfusundaki düşüş anlatılır. Otomobilin icadından önce şu kadar milyon at vardı; şimdi sadece şu kadar yüz bin at var diye... İnşallah başımıza öyle bir şey gelmez.

Bu konularla ilgilenen herkese hararetli bir şekilde -Türkçeye çevrilmedi galiba- şu kitabı tavsiye ediyorum: Weapons of Maths Destruction. Matematikli kitle imha silahları. Makine öğrenmesi ve “big data” denilen şeyin halihazırda -şu anda Amerika'dan birçok örnek veriliyor- nasıl fena sonuçlar verdiğine dair. Birkaç tane örneği var. Bir sürü karar mekanizmasında artık makine öğrenmesi ürünü algoritmalar kullanılıyor. Ve bunlar öğrenirken maalesef üstünde çalıştıkları verideki önyargıları da öğreniyorlar. Çünkü veriyi öğreniyorlar. Veri önyargı içeriyorsa onları öğreniyorlar. Az sonra örneğini göreceğiz. Bir de bunlar şeffaf olmadıkları için sorgulanamama özellikleri var. “Algoritma karar vermiş kardeşim”, “Böyle buyurdu algoritma”... Yani kimse onun ne demek istediğini ya da ne düşündüğünü de anlayamıyor ve insanlara zarar veriyorlar. Kitabı muhakkak okumanız lazım; birçok örnek var.

İngiltere'de hastaneyle ilgili meşhur bir örnek var. 1970'lerde, ilk makine öğrenmesini işe alma için kullanma denemelerinden biri. Hastaneye doktor olmak için başvuran insanların özgeçmişleri üzerinden bir eleme yapacak makine; üçte ikisini eleyecek. Neye göre eleyecek? “Geçmiş yıllarda kim başvurdu? Kim alındı? Kim alınmadı?” bilgisine göre. Makine onu öğreniyor. Geçmiş yıllarda eleme kriteri neyse ben de ona göre eliyorum, diyor. Sonra bakıyorlar ki işte Hindistan'dan Pakistan'dan başvuranları ötekilere oranla çok daha yüksek oranda eliyor. Kadınları erkeklere göre çok daha yüksek oranda eliyor. Neden? Çünkü geçmiş yıllarda öyle yapmışlar da ondan. Yani tıbbi yetenekten ziyade adam “madrass india”yı gördü mü -geçmişte o özelliğe sahip o kadar kişi elenmiş ki- pat, eleme kriteri budur diye eliyor. Bilgisayar da onu öğrenmiş. Sonra İngiliz hükümeti bunlara ceza veriyor. “Siz ırkçılık yapıyorsunuz, makine ırkçılık yapıyor” diye hastaneye ceza veriyor.

Sonra üniversite sıralamaları. Dergi, dünyanın en şahane üniversiteleri listesini basıyor ya... Derginin tek işi bu oldu zaten. Eminim farkındasınızdır üniversite mensupları!... Bunun da bir sürüsü var. Değişik üniversite sıralama endekslerinden birinde 1000. olmazsanız da diğer bir endekste 1000. olabilirsiniz. Öyle güzel bir şey var. Bunun yarattığı sıkıntılar. Bu adamlar, bu kriterleri hazırlarken önce rektörlere soruyorlar: “Sizce öteki üniversitelerin arasında bir harikalık sıralaması yaparsanız nasıl olur?” Bir şey çıkıyor. İlk sene ona göre sıralıyorlar üniversiteleri. Fakat herkes isyan ediyor tabii. Alt sıralarda kalanlar “Saçmalıyorsunuz. Tamamen kişisel görüşlere dayalı bu, veriye dayanmanız lazım,” diyorlar. Bunlar da veriye dayanmaya başlıyorlar. Veriye dayanırken de şöyle kriterler var: Bitirme oranı; yani “Kaç kişi giriyor? Onlardan kaç tanesi hakikaten mezun olup çıkabiliyor.” Ondan sonra “Kaç kişi başvuruyor? Onlardan kaç tanesi alınıyor?” Kaç kişiyi reddedebiliyorsa oran olarak kaç kişiyi reddedebiliyorsa o kadar harika bir üniversite demek ki falan... Böyle şeyler yapıyorlar. Buna göre bir sistem kuruyorlar. Ama mesela iyi üniversite için hani ücreti azsa ve iyi öğrenim veriyorsa o zaman o daha yüksek olması gerekir değil mi? Onu yapmıyorlar çünkü formüle önce onu koyarlarsa Harvard falan altta çıkıyor. Pahalı olduğu için. O yüzden onu koymuyorlar. Kendi kafalarındaki sıralamaya uygun olacak şekilde kullanacakları öznelilikleri seçiyorlar. Ve ondan sonra da Amerikan eğitim sistemi mahvoluyor. Çünkü herkes standart, kendisine dayatılmış olan bu formüle uymaya çalışıyor. Yani mesela bu sıralamada yukarı

çıkılmak için futbola yatırım yapmanız lazım. Neden? Hani onların futbol takımları falan oluyor ya. Böyle en pahalı futbol antrenörünü, çok akıllı bir öğrenci olmasa da iyi futbolcu öğrenciyi alıp - bunun sebebi ne? Çünkü o zaman size talep size çok artıyor... Böylece o demin dediğim kritere göre reddedebileceğiniz bir sürü başvuru geliyor. Böylelikle sıralamada yükseliyorsunuz. Ya da işte bitirme oranları... Mezuniyetten 9 ay sonra öğrencilerinin kaç tanesi işsiz filan diye bir şey var. Mesela hukuk fakültesinde mezuniyetten 9 ay sonra öğrencilerini kontrol ediyorlar. İşsiz falansa adama hemen böyle asgari ücretli, barda barmen olarak iş buluyorlar ki o şeyi geçsin. Muhakkak bunu öneririm yani bu kitapta böyle bir sürü örnek var...

Öğretmen değerlendirme sistemleri... New York'ta öğretmenleri belli bir şekilde, öğrencilerinin performansına, sınıflarının ortalamasına göre değerlendiriyorlar; altta kalanları işten atıyorlar. Kendisinden önceki öğretmen, aynı sınıfın önceki öğretmeni mesela notları şişirip vermişse, bu öğretmen hakkıyla vermişse, bu öğretmenin döneminde notlar düşmüş gibi oluyor ve bu öğretmen yok yere işten atılıyor. Öyle şeyler oluyor. Üstelik bu kararları hep makine verdiği için anlamsız şekilde kimse de buna itiraz edemiyor. Makinenin içini açtığınızda bu algoritmalar yine şirketlerin malı. Hem detayını göstermek istemiyorlar hem de onu kaç kişi anlayabilir ki... Böyle birçok örnek var. Hararetle tavsiye ederim kitabı.

Harari'nin kitaplarını ben çok beğeniyorum. Genel olarak insanlığın geleceği ile ilgili enteresan, düşünmeye değer fikirler ve olasılıklardan bahsediyor. Şimdi, Google her şeyimizi biliyor ya, diyelim bir hanımefendi bir beyefendiyle evlensem mi evlenmesem mi diye karar aşamasında. Diyor ki bunu Google senden daha iyi bilebilir; çünkü o senin de bütün özelliklerini biliyor, onun da bütün özelliklerini biliyor ya da birden fazla aday varsa onların da bütün özelliklerini biliyor. Ayrıca tıpkı sizin gibi olan ve geçmişte evlenmiş, evlenmemiş, onu değil de bunu seçmiş olan 68.000 başka insanın da özelliklerini biliyor ve o evliliklerin nasıl çıktığını da biliyor. Eğer bütün veriler böyle merkezi bir yerde varsa hakikaten o yüzde bilmem kaç olasılıkla sen aslında bununla evlenmek istiyorsun, ama böyle yaparsan beş sene sonra boşanacaksınız. Oysa bununla evlenirsen daha iyi olur falan gibi böyle tahminler yapabilir diyor. Özellikle bu yeni moda makine öğrenmesi ve sinirsel ağ sistemlerinin demin de dediğim özelliklerinden biri bunların ne öğrendiklerini, verdikleri kararları neden verdiklerini, onlara apaçık insan diliyle sorup yanıtını alamıyoruz.

Eski moda yapay zekâda bu olabiliyordu mesela. Benim tezimde bir tane Prolog programı çıkıyordu ortaya. Merak ediyorsanız bunu niye yaptı diye, açıp Prolog programı okuyup, “Hah şu yüzden yaptı” diye insan diline yakın yanıtlar alabiliyorduk. Oysa bir şey öğrenmiş olan bir sinir ağı. Buyurun sinir ağını size açtım, bakın dediğimizde milyonlarca, yüz milyonlarca sayıyla karşı karşıya kalıyorsunuz. O size, onun ne öğrendiği konusunda şeffaf bilgi vermiyor. Bunların ne düşündüğünü, ne bildiğini tam anlayamıyoruz. Kararlarını gerekçelendiremiyorlar. Ve eğer işlerinizi, demin anlattığım şekilde makine öğrenmesiyle öğrenilmiş kavramlara dayayarak yaparsanız veriden öğrenen sistemler doğal olarak geçmişte olan görüntü neyse onu öğreniyorlar. Kararlarınızı bundan sonra artık buna göre vermeye başlarsanız geçmişteki görüntüye siz hapsolmuş olursunuz.

Temel varsayımları, gelecekte de geçmişte ne olduysa gelecekte de aynı sistemin devam edeceği varsayımı üzerine. Oysa mesela insanlık öyle değil. “The Better Angels of Our Nature” adlı bir kitap var, Steven Pinker'ın onu da öneririm. Şiddet oranı, insanların birbirlerini öldürme oranı binlerce yıldır sürekli azalıyor. Hani ortaçağda çalıştırılmış böyle bir algoritmanın yapacağı tahminler çok daha karamsar olacaktı ve biz bütün ceza sistemimizi, hukuk sistemimizi ona kilitlemiş olsaydık, şimdiki, daha güzel olduğunu umduğumuz noktalara gelmemiş olacaktık.

Dikkat etmemiz lazım. Kötüye kullanımla ilgili aklıma gelen başka örnekler var. Bunu gördüğümde hoşuma gitmişti. Twitter mesajlarından insanların nerede olduğunu, o sırada ne içmekte olduğunu, alkol oranlarını öğrenen bir makine var. İçinde çeşitli anahtar sözcükler olan binlerce “tweet”i çekiyorlar, onları o Mekanik Türk denen insancıklara okutuyorlar. Bunlar hakikaten “ben sarhoşum” mu demek istiyor, yoksa alakasız mı? Şu anda içiyorum mu? Başka içen birisinden mi bahsediyor? Onları sınıflandırıyorlar, onların üzerinden makine öğrenmesini çalıştırıyorlar. Ve de “Kim nerede içiyor?” şeklinde bir harita çıkarabiliyorlar. Bunun hakikaten, bu teknolojiye sahip kişi veya kurumlar tarafından görülebilmesini istiyor musunuz? Twitter başka ne iş için kullanılabilir? Başka ülkelerde insanları otomatik birisi tutuklanacak veya 100 kişi tutuklanacaksa onların “tweet”lerinden seçip onları tutuklayalım işi için kullanılabilir. Allahtan bizde öyle bir durum yok görüldüğü gibi...

Bu sosyal ağların bir de şöyle korkunç kullanımı var. Çin, biliyorsunuz, duymuşsunuzdur; bir sosyal puanlama sistemi (“credit score”) gibi bir uygulamaya geçiyor. Çin’de bir arkadaşım var. Son zamanlarda konuşuyoruz, oradan da biliyorum. Artık sadece telefonunuzla çıkabilirsiniz, cüzdana gerek yok. Çünkü onların milli ve yerli WhatsApp’ları aynı zamanda ödeme de yapabiliyor. Hem rahatlık hem de bir yandan da siz ne ödediniz, ne aldınız bütün detayına kadar o sistem biliyor. Faturalarını zamanında ödedi mi? Sözlerine uydu mu? Sınavlarda kopya yaptı mı? Bütün bilgiler bir arada... Nerede oturuyor? Saatlerce bilgisayarda oyun mu oynuyor, yoksa bebek bezi mi alıyor? Ha demek ki sorumlu bir ebeveyn. Pozitif enerji paylaşıyor mu? Yani mesela ülke çok iyi gidiyor filan gibi sosyal medyada yazılar yazıyor mu, yazmıyor mu? Bunlar hakikaten “değerlendirmede kullandığımız kriterler” diye bu şirketin 2020’de herkese zorunlu olacak olan sosyal notlama sistemini kuran şirketin yetkililerinin söylediği şeyler. İnsanlar puanlar alıyor; 300 ile 1000 arasında mı ne puanlar alıyorlar. Puanınız yüksekse size daha çok kredi veriyorlar. Çeşitli ülkelere vizeyi daha kolay ya da sorunsuz alabiliyorsunuz. Eş bulma sitelerinde puanınız yükseliyor. Orada kadın sayısı, erkek sayısından daha az. Öyle bir sıkıntıları da var. Puanınız düşükse İnternetiniz yavaşlıyor, bazı restoranlara giremiyorsunuz. Vizeyi falan unutun zaten... Çeşitli işlere giremiyorsunuz. Böyle bir skorum sistemi kuruyorlar. 2020’de zorunlu olacak.

Bu haberi de muhakkak duymuşsunuzdur. Michal Kosinski diye bir psikolog var. Psikologların kullandığı beş boyutlu bir kişilik değerlendirme sistemleri var. Yani bir tür test yapıyorsunuz. O teste verdiğiniz yanıtlara göre sizin bu beş boyuttan hangisinde, ne tip bir insan olduğunuz çıkıyor. Ve $2^5 = 32$ değişik insan tipinden birine sizi oturtuyor. Anladığım kadarıyla yerleşik bir sistem. Hakikaten psikolojide bunun sağlam bir fikir olduğu konusunda hemfikirlik var. Kosinski, bunu Facebook’ta yapabileceğini fark ediyor. Çünkü Facebook’ta vakit geçirmek isteyen insanlar, “Hadi bakalım, sen nasıl bir insansın” falan konulu testleri doldurmayı da çok seviyorlar. Binlerce insana böyle doldurtuyor testleri. Böylelikle hem insanın tipini biliyor, hem de Facebook’a koyduğu öbür verileri biliyor. Buradan makine öğrenmesi yaparak, artık sadece Facebook çıktısından -yani neleri beğenip neleri beğenmediğinden, neye güldüğü, neye kızgın surat ikonunu bıraktığından falan- insanların zaten elinde olan diğer bilgileriyle bağlantı kurarak, sadece neyi “like” edip neyi “like” etmediğinden kişinin ırkını, cinsel yönelimini, hangi partiye meyil ettiğini, zekâ seviyesini, annesi babası boşanmış mı boşanmamış mıyı bakın sadece adamın neyi “like” edip neyi “like” etmediği kombinasyonundan tahmin edebilmeye başlıyor. Sonra da bir gün fark ediyor ki 10 “like” bilgisiyle kişiyi arkadaşlarından daha iyi tahmin edebiliyor. Daha sonra neyi “like” edeceğini veya belli bir ürünü alıp almayacağını, bir yeri tıklayıp tıklamayacağını... Böyle böyle 300 beğeniyle karısından/kocasından daha iyi tahmin edebilir hale geliyor. Daha çoğu ile adamın kendisinden daha iyi tahmin ediyor.

Bu teknolojiyi kullanan Cambridge Analytica diye bir şirket var. Bu şirket, Brexit oylamasında ve Trump'ın seçildiği seçimlerde kazanan tarafa danışmanlık veriyor. Bilin bakalım 2019'da kime danışmanlık verecek? Bu sistemin özelliği şu: İnsanları artık eski moda anketçiler gibi mağara devri tekniklerle değil, teker teker inceliyor yani. “Bu adam demokrat, ama bir yandan da zenci ve azıkcık da eşcinsel olduğu için Hillary'nin şu konuşmasını ona dinletirsek o gün sandığa gitmekten imtina edebilir” diye o adama sadece Hillary'nin o konuşmasını -biliyorsunuz Facebook'ta böyle hedefli şeyler gönderebiliyorsunuz, insanları bu şekilde seçip onlara özel yayın yapabiliyorsunuz-gösteriyorlar. Ya da insanlara birer birer adreslerine varıncaya kadar (bu işi yapan şirketin “web” sitesinden bu resim) işte şu adreste şöyle bir insan var, ona bunu dersin bize oy verebilir diye gezen siyasi parti görevlilerine haritalar, listeler verebiliyorsunuz.

Neyse içinizi kararttım, şimdi de yapay zekânın iyi sonuçları neler olabilir diyerek yavaş yavaş bitireyim. Bir kere yapay zekâ sayesinde insan troller işsiz kalabilir. Neden? Twitter'da falan üç beş kuruluş para karşılığında görevi -hedefli olarak, siyasi hedefler- genellikle insanlara hakaret etmek, onların mesajlarının altını doldurmak olan tipler var, bunlar yakında işsiz kalacak; çünkü bunun aynısını yapay zekâyâ yaptırmak çok kolay. Üstelik çok daha verimli şekilde. Tabii şimdi bu çok iyi olmadı galiba...

Biz bu işlere aslında çok temiz duygularla girmiştik. Yapay zekâ projesinin esas başlangıç hedeflerinden bir tanesi *insan zekâsı, yani zihin nasıl çalışır*, bunun benzerini yapmaya çalışarak ve nelerin zor, nelerin kolay olduğunu anlayarak çözmeye çalışmak. Umarız günün birinde burada edindiğimiz bilgiler insan zihninin nasıl çalıştığı, nasıl “hack”lenebildiği -ama iyi anlamda “hack”lenebileceği, neresine nasıl yaklaşmak gerektiği konusunda bize bilgiler verebilir. Zekâ genellikle ne kadar çok olursa o kadar iyi olan bir şey. Birçok uygulaması düşünülebilir. Mesela birkaç kez verdiğim bir örnek: Biz elektrik tasarrufu olsun diye yaz saatini kalıcı hale getirmek istiyoruz filan diye birisi bir kararname, bir kanun falan çıkarmaya çalışırsa, böyle şeylerde otomatik bir dijital asistan “Bir saniye, saçmalıyorsunuz, yaz saati zaten elektrik tasarrufu olsun diye...” diyen, böyle bir ekstra kontrol, uyarı sistemi üretmek için kullanılabilir.

Ciddi şekilde deminki Watson örneğinde gördüğümüz gibi çok fazla doktor var, çok fazla tıbbi deney var. İnsanların artık DNA bilgileri ve hangi DNA'yla ne çalışıyor bilgisi de konuluyor ama o makalelerinin hepsini okumak mümkün değil. O kombinasyonları üretip gerçekten günün birinde kişiye özel tedaviler yapay zekânın o büyük işlem gücü sayesinde, okuma gücü sayesinde çıkabilir. Bir de demin de dediğim gibi şu anda aklımıza gelmeyen iyi şeyler olabilir. Ufuk Çağlayan bunu hep anlatırdı. DARPA mıydı o? İnterneti icat eden adamlar, o sırada yazdıkları bir raporda, işte şöyle iyi şeyler olacak bilgisayarları birbirine bağlarsak diye anlatmışlar, ama e-posta akıllarına gelmemiş mesela. O yüzden bunda da aklımıza gelmeyen iyi şeyler çıkabilir. Böyle bir olasılık her zaman var.

“Bilgisayar Mühendisinin Sorumluluğu” başlığımız... Onunla bitireyim. Yapay zekânın karşıtı nedir? Doğal aptallık. Türkiye’de bilgisayarla da alakalı olarak dijital deliller denen zımbırtının kötüye kullanılmasıyla ilgili kötü bir dönem yaşadık. Hâlâ da yaşıyor olabiliriz aslında. O konuyla ilgili birkaç üniversiteden birkaç arkadaş, bazıları da burada şu anda, artık canımıza tak edince, biraz kelleyi koltuğa alıp “Sevgili mahkemeler, dijital belge, yani ben darbe yapacağım diyen bir Word dokümanı, o adamın darbe yapacağı anlamına gelmez. Çünkü o dokümanı başkaları yazmış olabilir,” gibi ilkokuldaki öğrencilerinin bile bildiği şeyleri adımızla sanımızla yazıp imzalamıştık. Raporda, o metinde imzası olan Coşkun Sönmez Hocamızı kaybettik. Onu da öteki hocalarımızla birlikte, bugün bu fırsat geldi, anmak istiyorum.

Çok sağ olun. Sorusu olan var mı? Bir soru var. Buyurun.

Dinleyici: Merhaba, IBM'de çalışıyorum. Bilgisayar mühendisiyim. ODTÜ, 92 yılı mezunuyum. Üniversiteler ya da bilgisayar mühendisliği bölümleri, yapay zekâ konusuyla ne kadar ilgili? Yapay zekânın dört bacağından birisi yetenek. Bir yandan bir şeyler konuşuyoruz. İşte endişeler var, iş kayıpları var siz de değindiniz ama bir yandan da yeni yeteneklere ihtiyaç var. Mesela en çok aranan pozisyonlarda biri veri bilimcisi, “data curation” (veri düzenleyicisi) gibi alanlar var. Düşünceleriniz nedir bu konuda?

Cem Say: Türkiye’de üniversitelerimiz, yapay zekâ konusunda -bilgisayar mühendisliği ya da biraz daha geneliyle övünmek gibi olsun- gayet iyi. Bizim hem akademik bakımdan hocalarımızın dünyadaki yayınları, araştırmaları açısından, hem öğrencilerimizin seçilme yöntemi zaten en iyi yöntem, bu açıdan hiç bir sıkıntımız yok. Bizim bölüm başkanımız burada, onun bir kriteri var “metrekareye düşen yayın sayısı” diye. Bölümün mekânı o kadar küçük ki orada dünya birincisi falanız. Ethem Alpaydın diye bir hocamız var, yapay öğrenme konusunda hakikaten dünya çapında bir isim. MIT Press'ten ders kitabı var; Türkçeye çevrildi. Yani övünmek gibi olsun. Zaten bu teknoloji de “Oraya gitmem lazım, yoksa öğrenemem” konusu olmayan bir teknoloji. Özellikle veri bilimi yavaş yavaş hakikaten “Evinizde öğrenebilirsiniz” tadında; birbirleriyle paylaşma biçimleri de var. Öğrenilebiliyor. Ama Türkiye'nin malum geri kalanında bir sürtünme faktörünün yüksek olması var. Öğrencilerin daha lisede dışarı gitmek istemesi var. Yani problem benim gördüğüm kadarıyla oralarda. Yoksa akademik olarak -nazar değmesin- gayet yerli ve milli başarıdan başarıya koşuyoruz...

Çok teşekkür ederim. Sağ olun...