

Aktivite 5

Yirmi Tahmin–*Bilgi teorisi*

Özet

1000 sayfalık bir kitapta ne kadar bilgi vardır? 1000 sayfalık telefon rehberinde daha mı çok bilgi vardır? Yüzüklerin Efendisi kitabında? Eğer bunu ölçebilirsek bu bilgiyi depolamak için ne kadarlık bir alana ihtiyacımız olduğunu tahmini olarak buluruz.. Örneğin, şu cümleyi okuyabiliyor musunuz:

B cımldk ssl hrflr kyptr.

Muhtemelen okuyabilirsiniz çünkü sesli harflerde fazla ‘bilgi’ yoktur. Bu aktivite bilgiyi ölçmenin bir yoluna giriş yapmaktayız.

Yetenekler

- 9 Sayıları karşılaştırmak ve bir sayı aralığında çalışmak
- 9 Genelden özele sonuç çıkarmak
- 9 Soru sormak

Yaşlar

- 9 10 ve üzeri

Malzemeler

- 9 İlk aktivite için malzeme gerekmez.
- 9 Uzatma aktivitesi için her çocukta şunların olması gerekir:
 - 9 Çalışma kağıdı aktivitesi: Karar ağaçları (sayfa 40)

Yirmi Tahmin

Tartışma

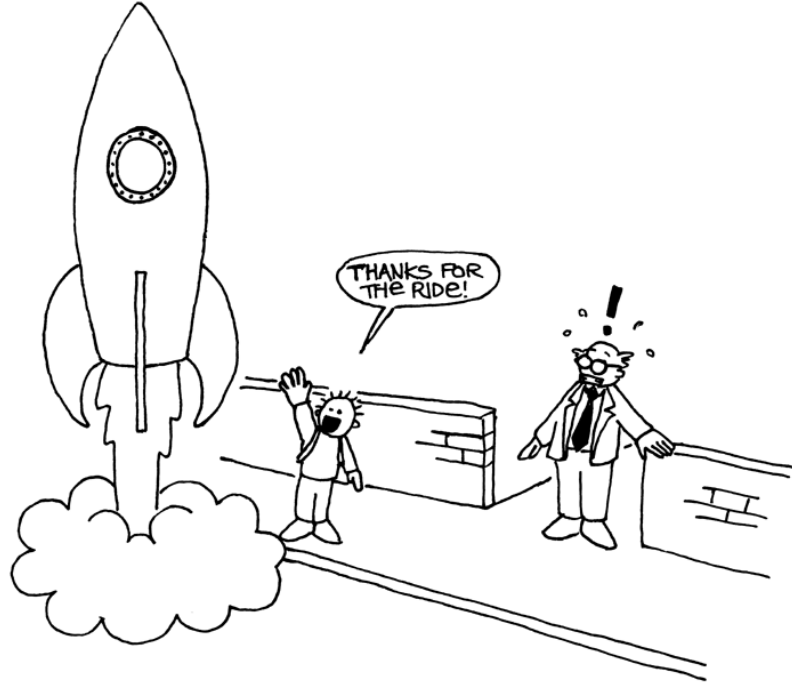
1. Çocuklarla bilginin ne olduğu hakkında tartışın.
2. Bir kitapta ne kadar bilgi olduğunu nasıl ölçebiliriz? Sayfa sayısı veya kelime sayısı önemli midir? Bir kitap bir diğerinden daha çok bilgiye sahip olabilir mi? Peki ya çok sıkıcı bir kitapsa veya çok ilginç bir kitapsa? “falan filan” kelimelerinden oluşan 400 sayfalık kitap aynı sayfa adedine sahip bir telefon defteri ile aynı bilgiye mi sahiptir?

Bilişim bilimcileri bilgiyi şu şekilde ölçer: Bir bilgi veya kitap ne kadar sürprizse o kadar çok bilgiye sahiptir. Zaten bildiğiniz bir şeyin size söylemesi pek bir şey ifade etmez—örneğin, her gün okula yürüyerek giden bir arkadaşınız size “Bugün okula yürüyerek gittim”—derse bir bilgi vermemiş olur, çünkü sürpriz bir şey değildir. Arkadaşınız, “Bugün helikopterle gittim” deseydi bu bir sürprizdi ve bilgi değeri yüksek olacaktı.

Peki bir mesajın sürpriz değeri nasıl ölçülebilir?

Bunun bir yolu bilgiyi ne kadar zor tahmin edilebileceğidir. Arkadaşınız, “Bil bakalım bugün okula nasıl gittim” dese ve yürümüş olsaydı bunu ilk tahminde bilebilirdiniz. Helikopterle gittiğini bulmak birkaç tahmin daha gerektirirdi. Uzay gemisi ile gittiğini bilmek ise çokça tahmin isterdi mutlaka!

Mesajların taşıdığı bilgi ne kadar sürpriz oldukları ya da olmadıkları ile alakalıdır. Birazdan okuyacağınız oyun bunun hakkında bize fikir verecek.



Yirmi Soru Aktivitesi

Bu 20 soru oyununun bir uyarlamasıdır. Seçilen bir çocuğa diğerleri sorular sorar ve yalnızca evet veya hayırdan oluşan cevaplar alır. Bu süreç tutulan sayı doğru bilindiğinde sona erer. Herhangi bir soru sorulabilir. Cevabın evet veya hayır olması şartı ile.

Öneriler:

Bir sayı tutuyorum:

- 9 1 ile 100 arasında
- 9 1 ile 1000 arasında
- 9 1 ile 1.000.000 arasında
- 9 herhangi bir tamsayı
- 9 6 numaradan oluşan bir dizi sayı (gruba uygun). İlkten sona kadar sırayla tahmin edin. (ör. 2, 4, 6, 8, 10)

Sorulan soru adedini sayın. Bu “bilgi”nin değerini hesaplamak için bir ölçüdür.

Devam eden Tartışma

Hangi stratejileri kullandınız? En iyileri hangileriydi?

1 ile 100 arasında bir rakamı tahmin etmek için her seferinde ikiye bölme yöntemiyle 7 adet tahmin gerektiğini söyleyelim. Örneğin:

50'den küçük mü?	Evet.
25'ten küçük mü?	Hayır.
37'den küçük mü?	Hayır.
43'ten küçük mü?	Evet.
40'tan küçük mü?	Hayır.
41'den küçük mü?	Hayır.
42 olmalı!	Evet!

İlginç bir şekilde aralık 1000'e çıkarıldığında gösterilmesi gereken çaba 10 kat artmıyor—yalnızca 3 adet daha fazla soru gerekiyor. Aralığın iki kat genişlemesi durumunda her seferinde bir adet daha fazla soru sorarız.

Uzatma: Bir mesajda ne kadar bilgi var?

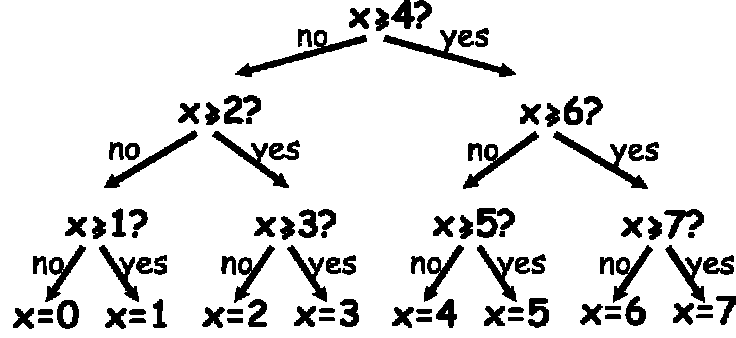
Bilişim bilimciler yalnızca sayı tahmini yapmazlar—aynı zamanda bir cümle ya da kelimedeki bir sonraki harfin ne olduğunu da iyi tahmin ederler.

Tahmin oyununu 4-6 kelimedeki oluşan kısa cümlelerle devam ettirmeye çalışın. Harfler doğru sırada (ilkinden sonuncusuna) tahmin edilmelidir. Birini harfler buldukça yazması için ayarlayın ve her harfte kaç tahmin yürütüldüğünü saydın. Evet veya hayır cevabı olabilecek her türlü soru kullanılabilir. Örnek sorular: “*t* mi?” “Sesli bir harf mi?” “*m* harfinden önce mi geliyor?” Her kelime arasındaki boşluk da aslında bir karakter olup harf gibi yine tahmin edilmelidir. Yani orada boşluk olduğunu tahmin etmelisiniz. Her cümlede hangi kısımların daha kolay tahmin edilebildiğini keşfetmeye çalışın.

Aktivite: Karar Ağaçları

Eğer her soruyu sorarkenki stratejiyi zaten biliyorsanız, bir mesajı iletirken soru sormadan iletebilirsiniz.

Buradaki çizelgede bir karar ağacı görüyorsunuz. Karar ağacında 0 ile 7 arasındaki bir sayıyı tahmin etmek için sorulması gereken sorular vardır:



5 numarasını 'tahmin' etmek için evet/hayır kararları sırayla nelerdir?

Herhangi bir numarayı çözmek için gerekli evet/hayır kararı sayısı kaçtır?

Şimdi çok büyüleyici bir şeye bakalım. 0, 1, 2 gibi en alt sıradaki sayıların altına ikilik düzende karşılıklarını yazın. (bkz. Aktivite 1).

Ağaca daha yakından bakın. Eğer hayır = 0 ve evet = 1 ise ne görüyorsunuz?

Numara tahmini oyununda soruları öyle seçeriz ki cevaplar numaraları aynen bu şekilde ikilik düzende ifade eder.

0 ve 15 arası sayıları tahmin etmek için gerekli olan karar ağacını tasarlayın.

Uzmanlara Ekstra: Birinin yaşını tahmin etmek için nasıl bir ağaç gerekir?
Bir cümlede bir sonraki harfi tahmin etmek için nasıl bir ağaç gerekir?

Tüm bunlar ne demek?

Calude Shannon ünlü bir Amerikan matematikçidir, aynı zamanda hokkabazlık yapar ve tek tekerli bisiklete biner. Bu tahmin oyunuyla ilgili bir çok incelemeler yapmıştır. Shannon, bilgi miktarını bitler olarak ölçer— her evet/hayır cevabı 1/0 bitine karşılık gelir. Bir mesajdaki “bilgi”nin miktarını halihazırda bildiğimiz bilgiye dayandırır.

Bazen öyle soru sorarsınız ki bir çok başka soruyu sormamıza artık gerek kalmaz. Böyle bir durumda mesajın bilgi içeriği düşüktür. Örneğin, her bozuk para fırlatıldığında bir bitlik bilgi oluşur: yazı veya tura. Fakat öyle bir bozuk para düşünün ki 10 atıştan 9’unda tura geliyor. Bu durumda ister inanın ister inanmayın bilgi bir bitten daha azdır. Bunu anlamak için şöyle bir deney yapalım. Basitçe şu soruyu soralım— “bundan sonraki *iki* atışta da tura mı gelecektir?” Örnekteki bozuk para $0.9 * 0.9$ ihtimalle yani 0.81 ihtimalle ikisinde de tura gelecektir yani %81 ihtimalle cevap evettir. %19 ihtimalle hayır cevabı gelir. %19 ihtimalle iki soru daha sormamız gerekir ki tam cevabı bilelim. Ortalamaya baktığımızda iki atışı bilmemiz gereken durumlarda %81 ihtimalle tek soru ile her iki atışı da tahmin edebiliriz! Yani %81 ihtimalle tek soru sorarak iki atışı da bulabiliyoruz!



Shannon bir mesajın bilgi içeriğine “entropi” demiştir. Entropi yalnızca olası sonuçların *adedine* bağlı değil—bozuk para atışında bu ikidir—aynı zamanda onun gerçekleşmesi *ihtimaline* bağlıdır. Mümkün görünmeyen olaylar, veya sürpriz bilgiler, çok daha fazla sayıda tahmine ihtiyaç duyar çünkü bize çok daha fazla bilgi verir—okula gitmek için helikopter kullanma durumundaki gibi.

Bilişim bilimcilere göre mesajın entropisi çok önemlidir. Bir mesajı entropisinden daha az yer kaplayacak şekilde sıkıştırmanız, ve en iyi sıkıştırma sistemleri bir tahmin oyununa eşittir. Sorulacak soruların listesi belli ise, bir bilgisayarın yalnızca cevapları tutması yeterli olur, daha sonra bilgiyi tekrar oluşturabiliriz! En iyi sıkıştırma sistemleri metin dosyalarını 10’da birine varan oranlarda sıkıştırabilir—bu oldukça büyük bir depolama alanı kazancı!

Bir kullanıcının yazı yazarken bir sonraki basacağı harfi tahmin etmek için de bu tahmin metodu kullanılabilir! Engelli insanlar için bu çok faydalı olur. Bilgisayar henüz onlar yazmadan bilgisayar önlerine seçenek sunarak yardımcı olur. Bir araştırmaya göre iyi bir yazılım iki veya daha fazla evet/hayır cevaplı soru sorarak bir sonraki harfi tahmin edebilir. Böylece mouse ve klavye kullanmada zorluk yaşayan biri kolayca yazı yazabilmektedir. Benzer bir sistem cep telefonlarında ‘yazmak’ için de kullanılır.

Solutions and hints

The answer to a single yes/no question corresponds to exactly one bit of information—whether it is a simple question like “Is it more than 50?” or a more complex one like “Is it between 20 and 60?”

In the number-guessing game, if the questions are chosen in a certain way, the sequence of answers is just the binary representation of the number. Three is 011 in binary and is represented by the answers “No, yes, yes” in the decision tree, which is the same if we write no for 0 and yes for 1.

A tree you would use for someone’s age might be biased towards smaller numbers.

The decision about the letters in a sentence might depend upon what the previous letter was.