

Prompt Tekniklerinin Model Performansına Etkisi: CoT ve CoT Olmayan Yaklaşımların Karşılaştırılması

Kart Oyunu Örneği Üzerinden Performans Karşılaştırması

Hazırlayan

İrem KUMLU

İçindekiler

1. Kullanılan Model ve Ortam	3
2. Problem Tanımı	3
3. Prompt Düzeni ve Yapısı	3
4. CoT ve CoT Olmayan Promptlar ve Yanıtları	3
5. Çıktıları Karşılaştırarak ve Farklılık Analizi	9
6. Yanıtların Doğruluğunu ve Netliğini Değerlendirilmesi	9
7. Chain of Thought Prompting Tekniğinin Model Performansına Etkisi	9

1.Kullanılan Model ve Ortam

Bu çalışmada, model olarak Claude 3.7 Sonnet kullanılmıştır. Denemeler ve çözümlemeler, Claude'un sağladığı sonuçlar üzerinden gerçekleştirilmiştir.

2. Problem Tanımı

Allemande adlı bir kart oyununda, dört oyuncunun her biri 52'lik standart bir destede 8 karttan oluşan bir ele sahiptir. Bir dizi atma yoluyla, oyuncular son ellerinin puan değerini en üst düzeye çıkarmaya çalışırlar. Renkler önemsizdir. As'tan 10'a kadar olan kartların puan değeri, kartlarının sayısının değeridir: örneğin, herhangi bir renkten beşli kart 5 puan değerindedir. Resimli kartlar (Vale, Kız ve Papaz) her biri 20 puan değerindedir. Charles en yüksek değerli son ele sahip midir?

Açıklama #1 : Charles'ın eli 117 puan değerindedir.

Açıklama #2 : Charles dışında hiçbir oyuncunun elinde dörtten fazla resimli kart yoktur.

3. Prompt Düzeni ve Yapısı

Raporda, Chain of Thought (CoT) ve CoT olmayan promptlar arasındaki farkları açıklarken, CoT promptlarının hiyerarşik bir yapı izlediği ve modele adım adım öğretmeyi amaçladığı dikkate alınmıştır. CoT promptlarında, her adımın belirli bir sırayla verilmesi, modelin daha mantıklı ve aşamalı bir çözüm üretmesini sağlar. Diğer yandan, CoT olmayan promptlar daha doğrudan ve tek adımlı sorular içerdiği için tek seferde üretmesi sağlanmıştır. Bu yapılar raporda sırasıyla şu şekilde düzenlendi:

1. CoT Olmayan Promptlar: Direkt çözüm önerileri ile sorular.
2. CoT Olunan Promptlar: Sorunun çözümünü adım adım açıklayan, her aşama için farklı detaylar sunan yapıda sorulmuştur.

4.CoT ve CoT Olmayan Propmtlar ve Yanıtları

CoT Olmayan Prompt :

“Allemande adlı bir kart oyununda, dört oyuncunun her biri 52'lik standart bir destede 8 karttan oluşan bir ele sahiptir. Bir dizi atma yoluyla, oyuncular son ellerinin puan değerini en üst düzeye çıkarmaya çalışırlar. Renkler önemsizdir. As'tan 10'a kadar olan kartların puan değeri, kartlarının sayısının değeridir: örneğin, herhangi bir renkten beşli kart 5 puan değerindedir. Resimli kartlar (Vale, Kız ve Papaz) her biri 20 puan değerindedir. Charles en yüksek değerli son ele sahip midir?”

Açıklama #1 : Charles'ın eli 117 puan değerindedir.

Açıklama #2 : Charles dışında hiçbir oyuncunun elinde dörtten fazla resimli kart yoktur.”

CoT Olmayan Prompt Yanıtı:

Bu problemi çözmek için Charles'ın elinin 117 puan değeri olduğunu ve diğer oyuncuların ellerinde dörtten fazla resimli kart olmadığını kullanacağım.

Öncelikle puanlamayı hatırlayalım:

- As'tan 10'a kadar kartlar kendi sayısal değerleri kadar puan (yani 1, 2, 3, ..., 10 puan)
- Vale, Kız ve Papaz her biri 20 puan

Charles'ın eli 117 puan değerinde. Bir el 8 karttan oluştuğundan, Charles'ın elinde olabilecek kartları düşünelim.

En yüksek puanı elde etmek için, Charles'ın elinde mümkün olduğunca çok resimli kart (Vale, Kız, Papaz) olması gerekir, çünkü bunlar en yüksek puan değerine sahiptir (her biri 20 puan).

Bir elde en fazla 8 kart olabileceğinden, Charles'ın elinde en fazla 8 resimli kart olabilir. 8 resimli kartın toplam değeri: $8 \times 20 = 160$ puan olurdu.

Ancak Charles'ın elinin değeri 117 puan. Bu durumda Charles'ın elinde bazı resimli kartlar ve bazı sayı kartları olmalıdır.

Charles'ın elinde olabilecek resimli kartların sayısını x olarak varsayalım. Bu durumda elinde $(8-x)$ tane sayı kartı olacaktır.

Resimli kartların toplam değeri: $x \times 20 = 20x$ puan Sayı kartlarının maksimum değeri: Eğer tüm sayı kartları 10 olursa, $(8-x) \times 10 = 80 - 10x$ puan

Tüm elin değeri = $20x + (\text{sayı kartlarının değeri}) = 117$ puan

Sayı kartlarının değeri = $117 - 20x$

Sayı kartlarının maksimum değeri = $80 - 10x$

Dolayısıyla: $117 - 20x \leq 80 - 10x$ Bu denklemi çözersek: $117 - 20x \leq 80 - 10x$ $117 - 80 \leq 20x - 10x$ $37 \leq 10x$ $x \geq 3.7$

Yani Charles'ın elinde en az 4 resimli kart olmalıdır.

Diğer oyuncuların her birinde en fazla 4 resimli kart olabileceği belirtilmiş. Toplam 12 resimli kart olduğundan (4 Vale, 4 Kız ve 4 Papaz), diğer üç oyuncunun elinde toplam en fazla 12 resimli kart olabilir.

Bu durumda, Charles'ın elinde resimli kartların sayısı, diğer herhangi bir oyunculakinden en az eşit veya daha fazla olacaktır. Ayrıca, Charles yüksek değerli sayı kartlarına (10, 9, 8 gibi) sahip olabilir.

Problem açıkça Charles'ın en yüksek değerli ele sahip olup olmadığını sorduğundan, diğer oyuncuların ellerinin en yüksek değerini hesaplayalım:

Bir oyuncunun elinde 4 resimli kart varsa (maksimum izin verilen), bu $4 \times 20 = 80$ puan eder. Kalan 4 kart için, en yüksek değerli kartlar 10, 9, 8 ve 7 olabilir, toplam 34 puan. Bu durumda bir oyuncunun alabileceği en yüksek puan $80 + 34 = 114$ puandır.

Charles'ın eli 117 puan değerinde olduğundan ve diğer bir oyuncunun eli en fazla 114 puan olabileceğinden, Charles en yüksek değerli ele sahiptir.

CoT Olan Prompt 1:

"Bir kart oyunu olan Allemande'de dört oyuncu vardır ve her biri 52'lik standart bir desteden 8 karttan oluşan bir ele sahiptir. Oyuncuların amacı, belirli kurallara göre ellerindeki kartların toplam puan değerini en üst düzeye çıkarmaktır. Puanlama şu şekildedir: * As'tan 10'a kadar olan kartlar, kartların sayı değerine eşittir. * Resimli kartlar (Vale, Kız, Papaz) her biri 20 puan değerindedir. Şimdi, verilen

bilgiler doğrultusunda Charles'ın elinin 117 puan olduğunu ve diğer oyuncuların hiçbirinin dörtten fazla resimli karta sahip olmadığını göz önünde bulundurarak en yüksek puanlı elin kimde olduğunu belirleyelim. Öncelikle, Charles'ın elindeki kartların dağılımını belirleyelim ve toplam puanı doğrulayalım."

CoT Olan Yanıt 1 :

Bu Allemande kart oyununda Charles'ın eli ve olası en yüksek puanlı eli analiz edelim.

Charles'ın elini belirlemek için önce puanlama sistemini inceleyelim:

- As'tan 10'a kadar kartlar kendi değerlerinde puan ($A=1, 2=2, \dots, 10=10$)
- Resimli kartlar (Vale, Kız, Papaz) her biri 20 puan

Charles'ın elinin 117 puan olduğunu biliyoruz. Olası kart dağılımlarını düşünelim:

Eğer Charles'ın elinde x adet resimli kart varsa, bunlar $20x$ puan değerindedir. Geriye kalan $(8-x)$ adet kart numaralı kartlardan oluşacaktır.

En fazla puan değeri için, numaralı kartların hepsinin 10 olmasını varsayalım:

- x adet resimli kart: $20x$ puan
- $(8-x)$ adet 10'luk kart: $10(8-x)$ puan
- Toplam: $20x + 10(8-x) = 20x + 80 - 10x = 10x + 80$

Bu toplam 117 puana eşit olmalı: $10x + 80 = 117 \Rightarrow 10x = 37 \Rightarrow x = 3.7$

Bu mümkün değil çünkü kart sayısı tam sayı olmalı. Demek ki tüm numaralı kartlar 10 değildir.

Deneme-yanılma yöntemiyle:

- 3 resimli kart (60 puan) + 5 numaralı kart (toplam 57 puan) = 117 puan

Numaralı kartların toplamı 57 olmalı. 5 kart için en yüksek toplam ($10+10+10+10+10=50$) bile 57'den küçük. Demek ki:

- 4 resimli kart (80 puan) + 4 numaralı kart (toplam 37 puan) = 117 puan

Numaralı kartların toplamı 37 olmalı. Bu makul bir senaryodur ve mümkün kart kombinasyonları vardır.

Diğer oyunculardan hiçbirisi 4'ten fazla resimli karta sahip olmadığına göre, en fazla 4 resimli kart olabilir:

- 4 resimli kart: 80 puan
- Geriye kalan 4 kart tümü 10 olursa: 40 puan
- Toplam maksimum: 120 puan

Sonuç olarak, bir oyuncunun elde edebileceği en yüksek puan 120'dir (4 resimli kart + 4 adet 10'luk kart). Charles'ın puanı 117 olduğuna göre, teorik maksimuma çok yakın ancak en yüksek puan değil. Başka bir oyuncunun 120 puan alabileceği bir el mümkündür.

Sonuç: Teorik olarak en yüksek puanlı el, 4 resimli kart ve 4 adet 10'luk kart içeren ve toplam 120 puan olan bir eldir. Charles'ın eli 117 puandır.

CoT Olan Propmt 2:

"Charles'ın eli toplam 117 puan olarak verilmiştir. Bunun doğru olup olmadığını doğrulamak için aşağıdaki gibi olası kart kombinasyonlarını inceleyelim: * Charles'ın elinde kaç resimli kart olabilir ve buna karşılık gelen puan değeri nedir? * Geriye kalan kartlar hangi sayısal değerlere sahip olabilir? Lütfen Charles'ın elindeki kartları tahmin ederek puanlamayı doğrula."

CoT Olan Yanıt 2 :

Charles'ın elinin 117 puan ettiğini doğrulamak için olası kart kombinasyonlarını inceleyeceğim.

Charles'ın elinde 8 kart var ve bunlardan bazıları resimli (Vale, Kız, Papaz) olabilir. Puanlama:

- Resimli kartlar (Vale, Kız, Papaz): Her biri 20 puan
- Sayısal kartlar (As'tan 10'a): Kendi değerleri kadar puan (1'den 10'a)

Charles'ın elinde olabilecek resimli kart sayılarını ve kalan puanı hesaplayalım:

1. 0 resimli kart (0 puan) + 8 sayısal kart (toplam 117 puan)
 - 8 sayısal kartın toplamı 117 olmalı
 - Bu imkansız çünkü en yüksek 8 sayısal kart (8 tane 10) = 80 puan eder
2. 1 resimli kart (20 puan) + 7 sayısal kart (toplam 97 puan)
 - 7 sayısal kartın toplamı 97 olmalı
 - Bu imkansız çünkü en yüksek 7 sayısal kart (7 tane 10) = 70 puan eder
3. 2 resimli kart (40 puan) + 6 sayısal kart (toplam 77 puan)
 - 6 sayısal kartın toplamı 77 olmalı
 - $77 \div 6 = 12.83...$ Bu imkansız çünkü sayısal kartlar en fazla 10 puan eder
4. 3 resimli kart (60 puan) + 5 sayısal kart (toplam 57 puan)
 - 5 sayısal kartın toplamı 57 olmalı
 - $57 \div 5 = 11.4...$ Bu imkansız çünkü sayısal kartlar en fazla 10 puan eder
5. 4 resimli kart (80 puan) + 4 sayısal kart (toplam 37 puan)
 - 4 sayısal kartın toplamı 37 olmalı
 - Bu mümkün! Örneğin: $10 + 10 + 10 + 7 = 37$ puan
 - Ya da: $10 + 10 + 9 + 8 = 37$ puan
 - Ya da: $10 + 9 + 9 + 9 = 37$ puan
6. 5 resimli kart (100 puan) + 3 sayısal kart (toplam 17 puan)
 - 3 sayısal kartın toplamı 17 olmalı
 - Bu mümkün! Örneğin: $10 + 6 + 1 = 17$ puan
 - Ya da: $9 + 7 + 1 = 17$ puan

- Ya da: $8 + 5 + 4 = 17$ puan

7. 6 resimli kart (120 puan)

- Bu durumda toplam 120 puan olur, bu 117'den fazladır, imkansız

Dolayısıyla, Charles'ın elinde iki olası kombinasyon vardır:

1. 4 resimli kart + 4 sayısal kart (toplam değeri 37)
2. 5 resimli kart + 3 sayısal kart (toplam değeri 17)

İki olası el:

- 4 resimli kart (80 puan) + 4 sayısal kart (10, 10, 10, 7) = 117 puan
- 5 resimli kart (100 puan) + 3 sayısal kart (10, 6, 1) = 117 puan

Charles'ın elinin 117 puan olduğu doğrulanmıştır.

Cot Olan Propmt 3:

*"Şimdi diğer oyuncuların ellerini inceleyelim. Bize verilen bilgiye göre hiçbir oyuncunun elinde dörtten fazla resimli kart yoktur. Buna dayanarak: * Bir oyuncunun elinde maksimum kaç resimli kart olabilir? * Geriye kalan kartlardan alabilecekleri en yüksek toplam puanı hesaplayalım. Bu hesaplamaları yaptıktan sonra Charles'ın en yüksek el değerine sahip olup olmadığını belirleyelim."*

Cot Olan Yanıt 3:

Bir oyuncunun elinde olabilecek maksimum resimli kart sayısını ve buna bağlı olarak elde edilebilecek en yüksek puanı hesaplayalım.

Verilen bilgilere göre:

- Hiçbir oyuncunun elinde 4'ten fazla resimli kart yok
- Her oyuncunun elinde 8 kart var
- Resimli kartlar (Vale, Kız, Papaz) 20'şer puan
- Sayısal kartlar kendi değerleri kadar puan (As=1, 2=2, ..., 10=10)

Maksimum puan için:

- Bir oyuncunun elinde maksimum 4 resimli kart olabilir ($4 \times 20 = 80$ puan)
- Kalan 4 kartın da mümkün olan en yüksek değerde olması gerekir, yani her biri 10'luk kart olmalı ($4 \times 10 = 40$ puan)
- Böylece elde edilebilecek maksimum toplam puan: $80 + 40 = 120$ puan

Charles'ın eli 117 puan olduğuna göre:

- $117 < 120$, yani Charles'ın eli teorik olarak mümkün olan en yüksek puana sahip değildir.
- Başka bir oyuncu 4 resimli kart ve 4 adet 10'luk kart ile 120 puan toplayabilir.

Dolayısıyla, Charles en yüksek puana sahip olmayabilir. Teorik olarak, 4 resimli kart ve 4 adet 10'luk karta sahip olan bir oyuncu 120 puan ile en yüksek puana sahip olacaktır.

Ancak bunun mümkün olup olmadığını kontrol etmek için 52 kartlık destede bulunan kart dağılımını düşünmemiz gerekir:

- Toplam 12 resimli kart vardır (her suitten 3 adet)
- Toplam 4 adet 10'luk kart vardır (her suitten 1 adet)

Eğer bir oyuncu 4 resimli kart ve 4 adet 10'luk kart almışsa, bu destede bulunan tüm 10'ları almış olur. Bu durumda bu oyuncunun 120 puan alması mümkündür ve Charles'ın elinden daha yüksek puana sahip olacaktır.

Sonuç: Charles'ın eli 117 puan ile yüksek bir puana sahiptir, ancak teorik olarak mümkün olan maksimum puan 120'dir. Dolayısıyla Charles en yüksek puana sahip olmayabilir.

CoT olan Prompt 4:

“Yaptığımız hesaplamalara dayanarak, Charles en yüksek puanlı ele sahip mi? Neden? Lütfen sonucu detaylıca açıkla”

CoT Olan Yanıt 4:

Hesaplamalarımıza göre Charles'ın elinde en yüksek puanlı el bulunmamaktadır. Bunu şu şekilde açıklayabilirim:

1. Charles'ın eli 117 puan değerindedir. Bu elinde ya:
 - 4 resimli kart (80 puan) + 4 sayısal kart (toplam 37 puan)
 - Veya 5 resimli kart (100 puan) + 3 sayısal kart (toplam 17 puan) olabilir.
2. Teorik olarak bir oyuncu maksimum 120 puan toplayabilir:
 - 4 resimli kart (80 puan) + 4 adet 10'luk kart (40 puan) = 120 puan
3. Verilen bilgilerde "hiçbir oyuncunun elinde 4'ten fazla resimli kart olmadığı" belirtilmiştir. Bu da demektir ki, bir oyuncunun elinde 4 resimli kart bulunabilir.
4. Standart bir 52 kartlık destede:
 - 12 resimli kart vardır (her suitten 3 adet: Vale, Kız, Papaz)
 - 4 adet 10'luk kart vardır (her suitten 1 adet)
5. Eğer bir oyuncu 4 resimli kart ve 4 adet 10'luk kart almışsa, ki bu kart dağılımıyla mümkündür, bu oyuncu 120 puan elde edecektir.

Sonuç olarak, Charles'ın eli 117 puan ile oldukça yüksek bir puan değerine sahip olmasına rağmen, teorik olarak mümkün olan maksimum puan 120'dir. Bu nedenle, Charles en yüksek puanlı ele sahip değildir. Başka bir oyuncu 4 resimli kart ve 4 adet 10'luk kart ile 120 puan toplayabilir ve bu Charles'ın puanından daha yüksektir.

5.Çıktıları Karşılaştırılarak ve Farklılık Analizi

Claude'un CoT (Chain of Thought) kullandığı yanıt ve CoT kullanmadan verdiği yanıt, çözüme ulaşma yollarında önemli farklar göstermektedir. CoT olmayan çözüm, problemi çözmek için daha doğrudan bir yaklaşım kullanır; burada, sadece verilen bilgilerin matematiksel olarak çözülmesi ve sonucunun elde edilmesi üzerinde durulmuştur. Charles'ın elinin 117 puan olduğu belirtilerek, basit bir denkleme dayalı olarak diğer oyuncuların elde edebileceği puanlarla karşılaştırma yapılır. Bu çözümde, tüm detaylar net bir şekilde göz önüne alınmamıştır ve sadece en olası senaryo üzerinden bir değerlendirme yapılmıştır. Ancak, CoT kullanılan çözüm, daha sistematik ve detaylı bir analiz sunar. Her adımda, hangi kartların hangi değeri taşıdığı, diğer oyuncuların ellerinin olasılıkları gibi faktörler dikkate alınarak sonuçlara varılır. CoT çözümünde, maksimum puanın 120 olabileceği belirtilmiş ve Charles'ın bu sınırın altında kaldığı vurgulanmıştır. Sonuç olarak, CoT kullanımı daha kapsamlı ve doğruluğu artıran bir yaklaşım sergilemiştir.

6.Yanıtların Doğruluğunu ve Netliğini Değerlendirilmesi

CoT olmayan yanıt, doğru hesaplamalar yapmasına rağmen, çözümdeki açıklık ve doğruluk bakımından bazı eksiklikler göstermektedir. Charles'ın eli 117 puan değeriyle verilmişken, diğer oyuncuların ellerinin detaylı bir şekilde analiz edilmemesi nedeniyle, elde edilen sonuç biraz daha yüzeysel kalmıştır. Sonuç, Charles'ın en yüksek puana sahip olduğu şeklinde verilmiş olsa da, bu çözümde diğer oyuncuların potansiyel en yüksek puanlarını göz ardı eden bir yaklaşım mevcuttur. Diğer oyuncuların kartlarının olasılıklarını hesaba katmadan yapılan bu çıkarım, doğruluk açısından daha düşük bir güven sağlar. CoT yanıtı ise çok daha kapsamlı bir analiz sunmuştur. Her adımda hesaplamalar yapılmış, teorik olarak ulaşılacak maksimum puan da dikkate alınmıştır. Bu da, Charles'ın 117 puanının aslında en yüksek puanlı elde edilebilecek puan olmadığını, başka bir oyuncunun 120 puan alabileceğini ortaya koymuştur. Netlik açısından, CoT yanıtı çok daha anlaşılır ve adım adım izlenebilirken, CoT olmayan yanıt daha fazla varsayıma dayanmakta ve eksik bilgi sunmaktadır.

7.Chain of Thought Prompting Tekniğinin Model Performansına Etkisi

CoT tekniği, problem çözme sürecinde büyük bir fark oluşturmuştur. CoT kullanılan yanıt, modelin sadece doğruluğunu artırmakla kalmamış, aynı zamanda sonuçları daha net ve anlaşılır bir şekilde sunmasına olanak tanımıştır. Her bir adımda, çözüm süreci açık bir şekilde takip edilerek modelin hangi mantıklı adımlarla sonuca ulaştığı gösterilmiştir. Bu sayede, Charles'ın en yüksek puanı alıp almadığı konusunda net bir sonuç elde edilmiştir.

CoT olmayan yaklaşım ise daha hızlı ancak yüzeysel bir çözüm sunmuştur. Burada, model yalnızca verilen bilgilerle hesaplamalar yapmış, ancak her olasılığı göz önünde bulundurmamıştır. Diğer oyuncuların kart kombinasyonları ve maksimum puan olasılıkları gibi faktörler dikkate alınmadığı için, modelin sonuçları daha eksik ve doğruluğu daha düşük olmuştur. Bu da CoT kullanımının modelin daha güvenilir ve sağlam sonuçlar üretmesini sağladığını ortaya koymaktadır.

Sonuç olarak, Chain of Thought prompting, modelin problem çözme sürecinde daha dikkatli ve mantıklı bir yaklaşım benimsemesine olanak tanır. Bu, modelin daha doğru ve güvenilir sonuçlar üretmesini sağlar. CoT tekniği, özellikle karmaşık problemlerde ve birçok olasılığın göz önünde bulundurulması gereken durumlarda büyük bir avantaj sağladığı sonucuna varılmıştır.