



"A Brain-Inspired Causal Reasoning Model Based on Spiking Neural Networks" Makalesinin İncelemesi

Giriş

Bu makale, sinirsel ağlar alanında **nedensel akıl yürütme** (causal reasoning) problemine biyolojik beyinden esinlenilen bir yaklaşım getirmektedir. **Beyin ilhamlı** (beyinden ilham alan) bir **nedensel akıl yürütme modeli**, geleneksel yapay sinir ağlarının (ANN) aksine, insan beynindeki sembolik bilgi temsilini ve çıkarım süreçlerini taklit eden **zirve ateşlemeli sinir ağları (Spiking Neural Networks – SNN)** kullanılarak geliştirilmiştir ¹ ². Araştırmacılar, günümüzde derin öğrenme tabanlı yapay sinir ağlarının sabit boyutlu vektörlerle bilgi temsil ettiğini ve bazı görevlerde başarılı olsa da insanın nedensel mantık yürütme biçiminden uzak ve **yorumlanabilirlik** açısından zayıf olduğunu vurgulamaktadır ³ ². Bu çalışma ise, SNN'lerin üçüncü nesil sinir ağı modeli olarak **daha biyolojik anlamlı** (daha biyolojik-plauzibl) olduğunu, ve **nedensel çıkarım, dizisel bilgi işleme, karar verme** gibi yüksek seviye bilişsel görevlerde de kullanılabileceğini ortaya koymaktadır ⁴. Makalenin amacı, **büyükanne hücresi hipotezi** (grandmother-cell hypothesis) gibi sinirbilim kavramlarından yararlanarak sembolik bilgi (örneğin "bir olayın bir diğerine sebep olması" ilişkisi) temsilini SNN'ler içerisinde gerçekleştirmek ve bu sayede **insan beyninin çıkarım süreçlerine benzer**, açıklanabilir ve yeni bilgi üretime imkân tanıyan bir yapay zeka modeli sunmaktadır ⁵ ⁶.

Özetle, makale beyin esinli mekanizmalar (ör. **nöron popülasyonu kodlaması** ve **ateşleme zamanına bağlı plastiçite, STDP**) ile bir SNN modeli tasarlayarak, **nedensel bilgi** olarak ifade edebileceğimiz **neden-sonuç ilişkili olaylar** arasındaki graf yapısını bu ağda temsil etmeyi ve bu yapı üzerinden **çıkarım (dedüksiyon ve endüksiyon)** yapabilmeyi hedeflemektedir. Bu yaklaşım, bir yandan biologik beyin modellerine dayandığı için sinirbilim araştırmalarına yeni bir hesaplamalı model sunmakta, diğer yandan yapay zeka alanında **daha sağlam, açıklanabilir** ve insan benzeri akıl yürütüebilen sistemlerin geliştirilmesi yönünde önemli bir adım olarak görülmektedir ⁷.

Modelin Temel Mantığı

Makalenin ortaya koyduğu modelde, **nedensel bir olaylar ağı** (yani bir **nedensel grafik**) SNN içinde birebir temsil edilmektedir. Bunu yapmak için **her bir olay (düğüm)** beyindeki belirli kavramların nöron gruplarında temsil edilmesi fikrinden ilhamla, **bir nöron popülasyonu** (yani çok sayıda nörondan oluşan bir alt-ağ) ile kodlanır ⁸. Bu, tek bir nöron yerine bir grup nöronun birlikte bir olayı temsil etmesi anlamına gelir ve sinirbilimdeki "grandmother cell" prensibi ile uyumludur; bir kavramın/hikayenin belli bir hücre grubu tarafından tanınması gibi ⁹ ¹⁰. Böylece model, tek bir nöron yerine bir popülasyon kullanarak **daha güvenilir ve gürültüye dayanıklı** bir temsil sağlar, ayrıca biyolojik beyin kodlamasına uygunluk (popülasyon kodlama, seyrek kodlama) kazanır ⁹ ⁶.

Bir **nedensel grafikteki ilişkiler (neden-sonuç bağıları)** ise SNN içindeki nöron popülasyonları arasında kurulan **sinaptik bağlantılarla** gerçekleştirilir ⁸. Yani eğer grafikte Olay-A, Olay-B'nin nedeni ise, modelde A'yı temsil eden nöron grubu ile B'yı temsil eden nöron grubu arasında uygun yönde kuvvetli bir sinaptik bağlantı oluşturulur. Bu bağlantı sayesinde, A grubu nöronlar ateşlendiğinde B

grubundakileri tetikleyebilecek ve böylece A olayı olduğunda B olayının gerçekleşmesini simüle eden bir sinirsel iletim sağlanacaktır. Makalenin getirdiği yenilik, bu bağlantıların **öğrenme yoluyla** kurulması ve ağın kendi kendine bu nedensel yapıyı edinmesidir. Özellikle, **Spike Timing Dependent Plasticity (STDP)** adı verilen biyolojik öğrenme kuralı kullanılarak, nöron popülasyonları arasındaki bağlantı ağırlıkları, ilgili olayların sıralı aktifleştirilmesiyle ayarlanır ⁸. STDP mekanizması kısaca, **ön-sinaptik nöronun bir ardıl-sinaptik nörondan** biraz önce ateşlemesi durumunda iki nöron arasındaki sinapsın kuvvetlendirilmesini, tersi durumda ise zayıflatılmasını öngören bir Hebb-vari öğrenme kuralıdır ¹¹. Modelde bu ilke, bir olayın nedeninin biraz önce, sonucun ise hemen sonrasında ateşlenmesi şeklinde uygulanarak, **nedensel yöndeki** sinapsların güçlenmesi sağlanmaktadır. Örneğin, A olayının B'ye neden olduğunu öğrenmek için eğitim esnasında A'nın nöron popülasyonu önce, B'nin nöron popülasyonu hemen ardından ateşlenir; bu kısa zaman farkı içinde STDP kuralı gereği A→B yönündeki sinaptik ağırlık artar (long-term potentiation, **LTP**), tersi yön zayıflar (long-term depression, **LTD**) ¹². Böylece eğitim sürecinin sonunda ağ içinde A'dan B'ye doğru güçlü bir bağlantı oluşur ve bu bağlantı, A'nın gerçekleşmesi durumunda B'yi aktive edebilecek şekilde ağı programları.

Şekil 1: Makalenin ele aldığı örnek nedensel grafiğin SNN ile temsilinin üç boyutlu şeması ¹³ ¹⁴. Bu senaryoda üstteki "Court Order" (mahkeme emri) olayı, "Captain" (kaptan) olayına yol açmakta; "Captain" olayı ise iki alt olay ("A" ve "B") tetiklemektedir. Son olarak bu iki olay birlikte "Death" (ölüm) olayının meydana gelmesine neden olmaktadır. Şemada her bir olay, daire içerisindeki **nöron popülasyonları** (çok sayıda küçük siyah nokta ile gösterilmiş) ile temsil edilmiştir. Oklar ise popülasyonlar arasındaki **sinaptik bağlantıları** (nedensel ilişkileri) göstermektedir. Eğitim süreci sonunda STDP kuralıyla bu ok yönlerinde güçlü bağlantılar oluşmaktadır ve ağ, **nedensel zinciri** biyolojik olarak gerçekleştirilebilir bir yapıya kavuşmaktadır.

Yukarıdaki şekildeki örnek senaryoda da görüldüğü üzere, model **statik bir nedensel grafiği** doğrudan doğruya bir SNN yapısına gömmektedir ¹⁵. Bunun için her bir düğüm (olay) belirli büyüklükte bir nöron kümesiyle temsil edilirken, aralarındaki kenarlar (nedensel-sonuç ilişkileri) bu kümeler arasındaki sinapslarla kodlanmaktadır. Eğitim aşamasında, grafikteki her nedensel ilişki için ilgili nöron popülasyonlarına **dışarıdan akım uyarımı** verilerek (yani ilgili popülasyonların kontrollü biçimde ateşlemesi sağlanarak) STDP yoluyla ağırlıklar ayarlanır ¹⁶. Makalede bahsedilen senaryoda, örneğin "**Court Order → Captain**" ilişkisini öğrenmek için "Court Order" popülasyonuna kısa bir akım enjeksiyonu yapılp hemen akabinde "Captain" popülasyonuna akım verilerek bu iki grup art arda ateşlettilir; böylece STDP mekanizması "Court Order"dan "Captain'a uzanan sinapsları kuvvetlendirir. Benzer şekilde "Captain → A", "Captain → B", "A → Death" ve "B → Death" ilişkileri de uygun ardışık uyarımlarla öğretilir. Sonuçta ağ içindeki sinaptik matriste, nedensel grafiğin yapısını yansitan güçlü bağlantılar açıkça ortaya çıkar. Nitekim çalışmanın deneySEL sonuçları, eğitim sonrası elde edilen sinaptik ağırlık dağılımının, yalnızca nedensel olarak bağlantılı olan popülasyonlar arasında yüksek ağırlıklar gösterdiğini doğrulamıştır ¹⁷. Bu da yöntemin, grafiği başarıyla öğrenebildiğinin bir kanıdır.

Teknik Detaylar: Nöron Modeli ve Öğrenme Kuralları

Modelde kullanılan bireysel nöronlar, biyolojik gerçekçiliği yüksek **sızan entegre ve ateşle (LIF – Leaky Integrate-and-Fire)** nöron modeliyle uygulanmıştır. LIF modeli, her bir nöronun zar potansiyelinin zamana bağlı olarak entegre edilip sızdırıldığı basit bir devre modeliyle tanımlanır. Denklem olarak, örneğin bir nöronun zar potansiyeli $V(t)$ için:

$$\$ \$ \tau \frac{dV}{dt} = -[V(t) - V_{\text{dinlenme}}] + I_{\text{girdi}}(t) \$ \$$$

biçiminde ifade edilebilir. Burada V_{dinlenme} nöronun dinlenim (istatistik) potansiyelini, $I_{\text{girdi}}(t)$ zamanla değişen giriş akımını (diğer nöronlardan sinaptik akımlar veya harici

uyarılar), τ ise membran zaman sabitini temsil eder. Nöronun potansiyeli, girişlerle belli bir **eşik değerine** yükseldiğinde nöron **ateşler** (bir **spike** üretir) ve potansiyeli bir reset değerine çekilir. Bu sayede her nöron, tipki gerçek bir biyolojik nöron gibi, girişlerinden aldığı uyarıları entegre edip belirli bir eşik aşıldığında çıktı ateşlemeleri üretir. Modelde bu mekanizma, zamana dayalı dizisel aktivite üretmek ve sinyallerin bir popülasyondan diğerine aktarılmasını sağlamak için kullanılır. Örneğin "Captain" popülasyonundaki nöronlar, "Court Order" popülasyonundan gelen girişler yeterince güçlü olduğunda hep birlikte eşik üstü potansiyellere ulaşarak **hepsinde hemen hemen eş zamanlı bir spike** patlaması yaratırlar. Bu da "Captain" olayın aktif hale geldiğini gösterir. Benzer şekilde, bu aktivite "A" ve "B" popülasyonlarına iletilerek onların ateşleme eşiklerini geçmesini sağlar vb. Bu süreç, ağı içinde **bir olaydan diğerine sinyal传递** şeklinde gerçekleşir ve bir zincirleme spike sekansı oluşturur.

Ağdaki öğrenme kuralı olarak vurgulanan **STDP (Spike Timing Dependent Plasticity)**, yukarıda genel mantığı verilen biyolojik bir Hebb öğrenmesi çeşididir. STDP'nin matematiksel formülasyonu, iki nöron arasındaki sinaps ağırlığındaki Δw değişiminin, **önce ateşleyen nöron** (presinaptik) ile **sonra ateşleyen nöron** (postsinaptik) arasındaki ateşleme zaman farkına (Δt) bağlı olacağını ifade eder. Tipik olarak bu bağımlılık şu şekilde modellenir:

- Eğer presinaptik nöron postsinaptikten **Δt kadar önce** ateşlediyse (yani $\Delta t = t_{\text{pre}} - t_{\text{post}} < 0$), sinaps ağırlığı **kuvvetlendirilir**. Kuvvetlendirme miktarı genellikle $A^+ \exp(-|\Delta t|/\tau^+)$ formunda verilir. Burada A^+ pozitif bir öğrenme katsayısı, τ^+ ise takviye (LTP) zaman sabitidir. Küçük zaman farkları (presinaptik hemen önce ateşlediyse) daha büyük güçlendirme, büyük farklar ise az güçlendirme sağlar ¹².
- Eğer presinaptik nöron postsinaptikten **sonra** ateşlediyse ($\Delta t > 0$), bu durumda sinaps **zayıflatılır**. Zayıflatma miktarı da benzer biçimde $A^- \exp(-|\Delta t|/\tau^-)$ formunda verilir, burada A^- negatif (ve mutlak değerce daha küçük) bir katsayıdır, τ^- ise depresyon (LTD) zaman sabitidir ¹².

Makalede kullanılan parametreler literatüre uygun şekilde seçilmiştir; örneğin STDP için tipik olarak $\tau^+ = \tau^- \approx 10$ ms, $A^+ > A^-$ ise LTP'nin LTD'den daha güçlü etki yapacağı şekilde ($A^+ > A^-$ alınmıştır (örneğin $A^+=0.5$, $A^-=-0.1$) ¹⁸). Bu parametre seçimi, presinaptik ateşlemenin postsinaptik ateşlemeden hemen önce gelmesi durumunda sinapsları belirgin ölçüde güçlendirirken, ters durumda ancak sınırlı bir zayıflama vermektedir. Böylece ağı, öğrenme sırasında sunulan nedensel ilişkilere odaklanarak ilgili bağlantıları kuvvetlice oluşturabilir.

Öğrenme sürecinde her bir nedensel bağlantı, **STDP zaman penceresi** denilen kısa bir zaman aralığında, ilgili nöron gruplarının sıralı şekilde uyarılmasıyla kurulur ¹⁶. "STDP penceresi" tipik olarak onlarca milisaniye mertebesindedir; bu süre içinde presinaptik grubun hemen ardından postsinaptik grubun uyarılması, aradaki sinapsları güçlendirmede etkilidir. Makalenin modelinde, harici bir tetikleme mekanizması ile hedeflenen nöron popülasyonlarına akım enjekte edilerek istenen sırada ateşlemeler sağlanmıştır ¹⁶. Örneğin eğitim aşamasında "Captain → A" bağlantısını kurmak için, "Captain" popülasyonuna akım verilerek (böylece "Captain" olayı simül edilerek) o popülasyondaki nöronlar ardisık bir spike treni üretir; hemen ardından "A" popülasyonuna akım verilir ve "A" popülasyonu ateşler. Bu ardisık ateşleme, "Captain"dan "A"ya giden sinapsları STDP ile güçlendirir. Bu işlem, grafikteki tüm kenarlar için (gerekirse birden fazla nöron grubunun aynı anda presinaptik olduğu durumlar da uygun şekilde zamanlanarak) tekrar edilir. Eğitim sonunda, ağıın sinaptik matrisi incelendiğinde, eğitilen nedensel grafik yapısının **izleri açıkça görülebilir**: ilgili popülasyonlar arasında sinaps ağırlıkları yüksek, alakasız popülasyonlar arasında ise düşük kalmıştır ¹⁷.

Ayrıca, her bir nöron popülasyonunun yapısı da biyolojik realizme katkı sunacak şekilde tasarlanmıştır. Popülasyonların içerisinde belirli oranda **uyarıcı (excitatory)** ve **bastırıcı (inhibitory)** nöronlar bulunmaktadır¹⁹. Bu sayede, hem popülasyon içi dinamiklerde denge sağlanmakta (aşırı uzun süreli kendiliğinden ateşlemeler önlenmekte) hem de biyolojik beyin dokularındaki E/I dengesi modele yansımaktadır. Dahası, harici olayların ağa beslenmesi için **Poisson kodlaması** gibi yöntemlerle girdi oluşturulmuştur; yani bir olayı temsil eden nöron grubuna dışarıdan belirli bir süre Poisson sürecyle dağıtılmış rastgele spike akışı verilerek, o olayın "oluştuğu" simüle edilmektedir²⁰. Bu rastgelelik, benzer olayların her tetiklenmesinde ağıda biraz farklı nöron alt kümelerinin ateşlemesini sağlayarak, modelin genelleme yapabilmesine ve gürbüzlüğüne yardımcı olmaktadır²⁰.

Nedensel Akıl Yürütme Süreci: Çıkarım ve Yorumlanabilirlik

Eğitim tamamlandıktan sonra, kurulan SNN modeli **nedensel grafiğe uygun çıkarımları** kendi dinamiğiyle yapabilmektedir. İki tür akıl yürütme senaryosu ele alınabilir: **dedüktif çıkarım** (ileri yönlü nedensel sonuç çıkarma) ve **endüktif çıkarım** (geri yönlü neden tahmini).

- **Dedüktif (İleri) Akıl Yürütme:** Bu senaryoda ağa belirli bir olayın gerçekleştiği bilgisi verildiğinde (yani ilgili nöron popülasyonu dışarıdan uyarılıp ateşlendiğinde), ağın sinaptik bağlantıları sayesinde **bu olayın nedensel sonuçları kendiliğinden ortaya çıkar**. Örneğin, eğitimli ağıda "Court Order" olayının gerçekleştiğini simüle edersek (üst popülasyona akım verip ateşletirsek), bu popülasyondan "Captain" popülasyonuna doğru güçlü sinapslar olduğu için "Captain" popülasyonu kısa bir gecikmeyle ateşlemeye başlar. Devamında "Captain" popülasyonundan "A" ve "B" popülasyonlarına sinyaller iletilir; uygun eşik aşıldığında bu iki popülasyon da spike üretir. Son olarak "A" ve "B"den "Death" popülasyonuna gelen birleşik girdiler sayesinde "Death" popülasyonu da aktif hale gelir. Böylece, sadece ilk nedeni besleyerek, model zincirleme bir etki ile nihai sonucu **kendiliğinden** üretmiştir. Bu, sistemin **nedensel grafiğe uygun çalışan bir benzetim** gerçekleştirdiğini gösterir. Nitekim makalede, eğitim sonrası yapılan testlerde, bazı ara düğümlere harici girdi verilmeden sadece başlangıç düğümü ve belki aradaki bir ilişki düğümü uyarıldığında bile, ağın eksik olan düğümü doğru biçimde "hatırlayabildiği" raporlanmıştır. Örneğin üçlü bir ilişki ağında (A-R-B şeklinde) sadece A ve R popülasyonlarına uyarı verildiğinde, B popülasyonunun da kısa süre sonra gücü şekilde ateşlemeye başladığı gözlemlenmiştir. Bu sonuç, SNN'nin öğrendiği bağlantılar sayesinde B'nin varlığını **çıkarabildiğini** (yani A ve R birlikte aktif olduğunda, aralarında ilişki olduğu için B'nin de olması gerektiğini "düşündüğünü") göstermektedir. Bizim incelediğimiz nedensel örneğe dönersek, sadece "Court Order" olayı verildiğinde modelin bir süre sonra "Death" olayını ateşlediği görülür; aradaki tüm zincir (Captain, A, B) ağı içi dinamikle aktifleştirilmiştir. Bu şekilde model, **nedensel çıkarımı** ileri doğru başarabilmektedir⁸.
- **Endüktif (Geri) Akıl Yürütme:** Bir diğer önemli yetenek, ortaya çıkan bir sonucun olası nedenlerini belirleyebilmektir. Geleneksel yapay sinir ağları genelde ileri yönlü hesaplama yaptığı için, bir çıktıının geriye dönük sebeplerini şeffaf şekilde vermekte zorlanırlar (bu tip ağlar "kara kutu" gibi davranışır). Oysa bu beyin-esinli model, **oluşan spike desenlerinin analiziyle** geriye dönük çıkarıma olanak tanır²¹. Ağ içinde belirli bir popülasyonun (örneğin "Death" popülasyonunun) ateşlediğini gözlemediğimizde, bu popülasyona güçlü giriş yapan hangi popülasyonların kısa süre önce ateşledidine bakarak sonucun muhtemel nedenlerini saptayabiliyoruz. Makalede bu amaçla **harici bir değerlendirme fonksiyonu** tanımlanmıştır²¹. Bu fonksiyon, belirli bir zaman anındaki ağın ateşleme durumunu, hafızada depolamış nedensel kalıplarla karşılaştırarak en uyumlu olanı belirler²² ²³. Basitçe ifade etmek gerekirse, ağın o andaki etkin nöron seti, hafızadaki "hangi olaylar dizisine" karşılık gelmektedir sorusunu cevaplamaktadır. Örneğin "Death" popülasyonunun ateşlemesi gözlendiğinde, eğer hemen öncesinde "A" ve "B" popülasyonlarının spike ürettiğini ve onların da öncesinde "Captain"

popülasyonunun etkinliğini kayıt altına alıysak, “**Death neden oldu?**” sorusuna doğal olarak “**Captain olayının tetiklediği A ve B olayları nedeniyle**” yanıtını verebiliriz. Bu zincir, en başta “Court Order” popülasyonun da ateşleyip ateşlemediğine bakılarak daha da geriye götürülebilir. Makalede bahsedilen değerlendirme fonksiyonu tam da bu tür bir **neden izleme (reasoning path tracing)** işlemine olanak sağlamakta ve her bir çıkışın arkasındaki zinciri görünür kılmaktadır²¹. Bu sayede model, klasik derin ağların aksine **şeffaf ve açıklanabilir** bir akıl yürütme süreci sunmaktadır. Yazarlar, CRSNN modelinin bu özelliğinin, aynı görevi yapan geleneksel ANN modellerine kıyasla çok daha yorumlanabilir olduğunu vurgulamışlardır²¹. Bu durum, özellikle güvenlik-kritik uygulamalarda (örn. otonom araçlarda kararın hangi sebeple alındığının anlaşılması gibi) ve bilimsel keşif süreçlerinde (örn. bir hipotezin hangi sebeple çıkarıldığıının izlenmesi) son derece değerli bir katkıdır.

Deneysel Kurulum ve Sonuçlar

Makaledeki deneysel çalışma, yukarıda açıklanan modelin küçük ölçekli bir senaryoda doğrulanmasını içermektedir. Yazarlar, gerçek dünyadan esinlenilmiş basit bir **nedensel senaryo** tanımlamışlardır (Şekil 1'de gösterilen örneğe benzer). Bu senaryoda 5 adet olay düğümü ve aralarında nedensel ilişkiler bulunmaktadır: Örneğin **Court Order** (Mahkeme Emri) verildiğinde bir **Captain** (Kaptan) olayı meydana gelmekte; **Captain** gerçekleştiğinde iki farklı alt olay (**A** ve **B**) tetiklenmekte; **A** ve **B** birlikte olduğunda ise bir **Death** (Ölüm) olayı ortaya çıkmaktadır. Bu yapıda **dallanma** (bir nedenin birden çok sonuca yol açması) ve **birleşme** (bir sonucun birden fazla nedene bağlı olması) durumları bir arada bulunmaktadır. Araştırmacılar, bu grafiği SNN modeline dönüştürüp başarıyla eğittikten sonra, modelin **bu zinciri doğru şekilde öğrenip öğrenmediğini ve çıkışım yapabildiğini** test etmişlerdir.

Deneysel kurulumda her bir olay için belirli büyülüklükte nöron popülasyonları alınmıştır (örneğin her düğüm için 100 civarı nöron kullanıldığı varsayılabılır; makalede net sayı belirtilmişse de burada amaç kavramsal büyülüklüğü anlamaktır). Başlangıçta ağıdaki tüm sinaps ağırlıkları küçük rastgele değerlerle başlatılmıştır. Ardından **eğitim aşaması** olarak, her nedensel bağlantı için STDP kuralına uygun ardışık uyarımlar gerçekleştirilmiştir. Örneğin “Court Order → Captain” bağlantısını güçlendirmek için defalarca “Court Order” popülasyonu hemen takiben “Captain” popülasyonu uyarılmıştır. Aynı şekilde “Captain → A”, “Captain → B”, “A → Death” ve “B → Death” için de tekrar eden ardışık ateşlemeler uygulanmıştır. Her böyle ikili ya da üçlü ardışık ateşleme sekansı, ilgili sinapslarda kademeli bir ağırlık artışı sağlamıştır. Yeterli sayıda tekrar sonrası ağın sinaptik bağlantıları, nedensel grafiğin topolojisini yansıtır hale gelmiştir. Eğitim süresinin toplamı ve iterasyon sayıları, ağ boyutuna göre değişmekle birlikte, bu senaryo için milisaniyeler mertebesinde birden çok timeslot (örneğin 100 ms'lık pencereler halinde) periyotta uyarımlar verilerek gerçekleştirilmiştir. Örneğin 0–100 ms arası Court Order & Captain, 200–300 ms arası Captain & A, 300–400 ms arası Captain & B, 400–500 ms arası A & Death, 500–600 ms arası B & Death uyarılmış olabilir (bu sıralama temsilidir). Her eğitim döngüsü sonrası, ağın ilgili sinapslarının kuvvetlendiği, yazarlarca sunulan şekillerde sinaps ağırlık matrisi görselleriyle gösterilmiştir. **Şekil 2**, eğitim sonundaki sinaptik ağırlık dağılımının bir örneğini (farklı bir senaryo olan üçlü A-R-B bilsisinin öğrenilmesi için) göstermektedir. Bu görselde, sadece ilgili nöron popülasyonları arasında (A ile R, R ile B arasında) yüksek ağırlıklar oluştuğu, popülasyonların kendi içlerinde de zayıf bağlantıların var olabildiği görülmektedir²⁴ ¹⁷. Bu, STDP öğrenmesinin **hedeflenen yapıyı başarıyla yakaladığını** doğrular niteliktedir.

Test aşamasında ise modele nedensel zincirin başlangıç noktasından tetiklemeler verilmiş ve sonucun kendiliğinden ortaya çıkıp çıkmadığına bakılmıştır. İlk senaryo olarak, **dedüksiyon testinde**, grafiğin kök nedeni sayılan olaya harici uyarı verilmiştir (ör. Court Order). Bunun sonucunda, yukarıda anlatıldığı gibi, ağ içi dinamikle sırasıyla diğer tüm ara düğümler ve sonunda hedef düğüm (Death) ateşleme üretmiştir. Yazarlar, ağın tetikleme dizisini zamana bağlı grafiklerle sunmuşlardır. Örneğin Court Order verildiğinde

ilk 10 ms içinde Captain'in spike oranında artış, 20–30 ms civarında A ve B'nin ateşlemeye başlaması, 40 ms civarında ise Death popülasyonunun yükselmesi gibi bir zaman diyagramı ortaya konulmuştur. Bu, modelin **nedensel akışı doğru zamanda** yansittığını gösterir. Ayrıca **kısmi ipucu ile çıkarım** da test edilmiştir: Örneğin orta kademedeki bir olay tetiklenmeden (veya zayıf tetiklenerek) sadece kısmi girdilerle final çıktıının oluşup oluşmadığı incelenmiştir. Bu sayede modelin, eksik bilgi durumunda bile belleğindeki grafiği kullanarak eksik parçayı doldurup dolduramadığı gözlemlenmiştir. Üçlü ilişki testinde örneğin A ve R verildiğinde B'nin de canlandığı (yani modelin B'yi "hatırladığı") rapor edilmiştir. Benzer şekilde, tam zincirde belki sadece "Court Order" ve onun bir alt dalı tetiklenip, diğer dalın ağ tarafından uyarılıp uyarılmadığı test edilmiş olabilir. Genel olarak sonuçlar, **CRSNN modelinin hem ileri hem geri yönde nedensel çıkarımları isabetli şekilde gerçekleştirebildiğini** ortaya koymuştur⁸. Bunun yanında, yazarlar modelin performansını klasik bir derin öğrenme modelinin performansıyla sayısal olarak karşılaştırmamış olsa da (IJCNN makalesi bir kavram kanıtlama çalışması olduğundan), daha sonraki genişletilmiş çalışmalarında SNN tabanlı yaklaşımın bir grafik yapısındaki çıkarım görevlerinde yapay sinir ağları (ör. Grafik Konvolüsyon Ağları, GCN) ile **benzer doğruluğa daha hızlı ve biyolojik anlamlı şekilde** ulaşabildiğini göstermişlerdir^{2 25}. Bu, modelin sadece biyolojik olarak ilginç değil, aynı zamanda yapay zeka uygulamaları için de **verimli ve rekabetçi** olabileceğine işaret etmektedir.

Son olarak, modelin **yorumlanabilirlik avantajı** testlerde niteliksel olarak ortaya konmuştur. Ağın çarptığı sonuçlarda, her bir sonucun hangi girdilerden dolayı tetiklendiği, harici değerlendirme fonksiyonuyla veya basitçe spike zamanlamalarını izleyerek anlaşılabilmektedir. Örneğin Death çıktısı alındığında, sistem kullanıcısı ya da geliştiricisi hangi popülasyonların önce ateşlediğini (A ve B'nin mi, yoksa belki farklı bir yolun mu) görerek karar sürecini takip edebilir. Yazarlar, bu sayede klasik yapay zeka sistemlerine kıyasla **şeffaf bir neden-sonuç silsilesi** sunulduğunu ve modelin **İç mantığının denetlenebilir olduğunu** belirtmiştir²¹.

Uygulama Alanları ve Potansiyel Etkiler

Bu beyin esinli nedensel akıl yürütme modeli, hem nörobilimsel araştırmalar hem de yapay zeka uygulamaları açısından heyecan verici potansiyele sahiptir.

1. Yapay Zeka ve Bilişsel Sistemler: Günüümüz yapay zeka sistemlerinde **sağduyu bilgisi ve nedensel çıkarım** en zorlayıcı konulardandır. Bu model, ConceptNet gibi büyük ölçekli bilgi grafiklerindeki kavramları ve aralarındaki ilişkileri SNN yapısında temsil etmenin yolunu açmıştır^{26 27}. Daha sonraki bir çalışmada yazarlar, bu yaklaşımı milyonlarca düğüm ve ilişki içeren bir bilgi ağı üzerinde uygulayarak, SNN'nin **ortak akıl (commonsense) bilgilerini** öğrenip sorulara cevap verebildiğini göstermişlerdir²⁸
²⁹. Dolayısıyla model, akıllı asistanlar, robotik sistemler veya karar destek mekanizmaları gibi alanlarda, bilgi tabanlı çıkışım yapma görevlerinde kullanılabilir. **En büyük avantajı**, bu çıkarımların **insan tarafından takip edilebilir mantık zincirleri** ile gerçekleşmesi, böylece sistemin **açıklanabilir yapay zeka (XAI)** kriterlerine uygun olmasıdır. Örneğin, otonom bir robot bu modelle donatılırsa, bir tehlikeli durumu (örneğin "yangın alarmı çalışıyor") algıladığında bunun muhtemel nedenlerini (örneğin "duman sensörü tetiklendi, mutfakta yangın olabilir") çıkarabilir ve bu zinciri kullanıcıya açıklayabilir.

2. Robotik ve Otonom Sistemler: Modelin sunduğu biyolojik plausibilite, **robotik uygulamalarda yeni bilişsel yetenekler** geliştirmede kullanılabilir. Nitekim, Fang ve Zeng'in bu çalışmasından ilhamla geliştirilen "**beyin esinli robot acı modeli**" buna güzel bir örnektir^{30 31}. Bu modelde bir robot, kendi vücut durumunu ve algılarını yine SNN popülasyonları olarak temsil etmekte; **STDP ile öğrenilmiş nedensel bağlantılar** sayesinde, vücudunda hasar olduğunda "acı" çıktısını üretmektedir^{32 33}. Örneğin robot kolunun durumu (düz veya büyük olması) ile sensör okumaları arasındaki ilişkiyi SNN öğrenmekte, beklenen durumdan sapma olduğunda hata ve acı nöron popülasyonları ateşleyerek

robotu uyarabilmektedir³⁰ ³¹. Bu tür bir uygulama, modelimizin robotik öz-farkındalık, öngörülebilir arıza tespiti, ve güvenlik davranışları (tehlikeden kaçınma gibi) kazanmasında rol oynayabileceğini göstermektedir. Yani CRSNN prensipleriyle, bir robotun kendi deneyimlerinden **nedensel öğrenme yapıp** gelecekteki olayları tahmin etmesi, hatta **olasılık dahilinde oluşacak istenmeyen durumları önceden sezip önlem alması** mümkün hale gelir.

3. Sinirbilim ve Biyolojik Modelleme: Makale sadece mühendislik açısından değil, aynı zamanda beynin bilişsel işleyişini anlama açısından da değer taşır. İnsan beyninde kavramların ve kavramlar arası ilişkilerin tam olarak nasıl temsil edildiği ve işleme alındığı hala araştırma konusudur. Bu çalışma, **bellekteki kavramların bir grup nöronla temsil edilmesi (grandmother cell) ve bu gruplar arası sinaptik bağlantıların bilgi grafikleri oluşturacak şekilde düzenlenmesi** fikrini sayısal olarak sınımsızdır. Böylelikle, nörobilimcilere beynin **ortak-akıl bilgisini** nasıl kodlayabileceğine dair test edilebilir bir model sunulmuştur³⁴. Makalenin yazarları, çalışmalarının **beynin sağıduyu bilgiyi temsil etme biçimini araştırmak için bir temel oluşturduğunu** belirtmektedir⁷. Özellikle prefrontal korteks ve medial temporal lob gibi alanların sembolik/bilişsel içerikleri nasıl işlediğini anlamak üzere bu tip modeller kullanılabilir. Gelecekte, deneysel sinirbilim verileri (örn. fMRI veya elektrofizyoloji ile elde edilen kavram haritaları) ile bu model karşılaştırılarak, beynin gerçek ağ yapısı ile modelin çıktıları arasında benzerlikler olup olmadığı incelenebilir.

4. Enerji-verimli Hesaplama ve Donanım: SNN'ler, klasik yapay sinir ağlarına kıyasla **olay tabanlı ve seyrek ateşlemeli** yapıları sayesinde çok daha az enerji tüketimiyle çalışabilmektedir. Bu nedenle, modelin donanımsal karşılığı olarak **nöromorfik çipler** üzerinde uygulanması mümkündür. Intel Loihi veya IBM TrueNorth gibi nöromorfik işlemciler, SNN prensipleriyle çalıştığından, nedensel akıl yürütme modellerini gerçek zamanda ve düşük güçle çalıştırabilecek platformlar sunar. Bu sayede, örneğin uç bir IoT cihazında, buluta ihtiyaç duymadan gerçek zamanlı nedensel çıkarım yapılabilir. Bir endüstriyel makinenin arıza nedenini yerinde tespit eden akıllı sensörler, veya bir cep telefonunda kullanıcı davranışlarından sebep-sonuç ilişkileri öğrenen asistan uygulamaları, bu teknolojinin potansiyel kullanım alanlarındandır.

5. Yazılım ve Bilgi İşlemde Yorumlanabilirlik: Günüümüz derin öğrenme modellerinin en büyük handikaplarından biri, sonuç üretirken kullandıkları iç mantığın kullanıcı tarafından anlaşılamamasıdır. Bu çalışma, kompleks grafik ilişkileri dahi barındırsa, ağı içi aktivite paternlerinin anlamlanırılabildiği bir sistem önermektedir. Bu, özellikle hukuk, tip, finans gibi **kararların kritik olduğu** ve açıklanması gereken alanlarda (örneğin bir yapay zekanın bir teşhisini neye dayanarak koyduğu, bir kredi başvurusunu neden reddettiği vb.) büyük önem taşır. CRSNN gibi bir model, kararın ardından nedensel zinciri açıkça sunabildiği için, güven ve şeffaflık sağlar. Örneğin makine öğrenimi tabanlı bir tıbbi tanı sisteme bu tarz bir nedensel SNN entegre edildiğinde, sistem sadece "hastada X sendromu var" demekle kalmayıp, "çünkü bulgular A ve B tespit edildi, bu da genellikle C nedeniyle olur, C de X sendromunun bilinen bir nedenidir" şeklinde izah sunabilir. Bu tip açıklamalar, yapay zekaya olan güveni artıracak ve **etik yapay zeka** prensiplerine uygunluğu sağlayacaktır.

Sonuç ve Değerlendirme

Sonuç olarak, "A Brain-Inspired Causal Reasoning Model Based on Spiking Neural Networks" makalesi, yapay zeka ile sinirbilim kesiminde yenilikçi bir katkı sunmaktadır. Model, **beyindeki sembolik akıl yürütme süreçlerini dijital bir SNN ortamına taşıyarak** nedensel ilişkileri öğrenebilen ve bu ilişkiler üzerinden çıkarım yapabilen bir çerçeve ortaya koymustur. Teknik açıdan, nöron popülasyonu tabanlı kavram temsili, STDP ile kendini organize eden grafik yapısı ve harici değerlendirme fonksiyonıyla yorumlanabilirlik gibi unsurlar bir araya getirilerek özgün bir mimari oluşturulmuştur. Deneysel sonuçlar, modelin basit senaryolarda beklendiği gibi çalıştığını, yani nedensel grafiği başarıyla

ıçselleştirip doğru ileri-geri akıl yürütmemeyi yapabildiğini göstermektedir. Bu çalışma, ileride **daha karmaşık bilgi ağlarına** (örneğin tam teşekkülü bilgi grafikleri, sosyal ağlarda nedensel etki analizleri v.b.) uygulanabilecek, ölçeklenebilir ve biyolojik temelli bir yöntem için ilk adımı temsil etmektedir.

Makalenin en önemli katkılarından biri, **yorumlanabilir ve beyine yakın AI sistemleri** geliştirme vizyonudur. Yazarlar, klasik derin öğrenme yaklaşımının aksine, önerdikleri modelin her adımının izlenebilir olduğunu ve bunun hem mühendislikte hem de bilimde büyük bir avantaj olduğunu dile getirmiştirlerdir ²¹. Nitekim devam eden çalışmalarında bu doğrultuda ilerlenerek daha büyük ölçekli SNN bilgi-ağı modelleri kurulmuş ve bunların performans açısından da rekabetçi olabildiği gösterilmiştir ²⁵. Elbette, bu modelin de geliştirmeye açık yönleri bulunmaktadır. Örneğin, mevcut haliyle statik bir grafiği öğrenmektedir; gelecekte **dinamik olarak değişebilen nedensel ilişkileri** (zamanla yeni ilişkiler öğrenme veya mevcut ilişkileri unutma gibi) uyarlayabilen genişletilmiş kurallar (ör. **ödül modülatonlu STDP** olarak bilinen R-STDP mekanizmaları ³⁵ ³⁶) eklenebilir. Yine de, bu çalışma bilim camiasına **"beyin gibi düşünən"** sinir ağları tasarılama yolunda somut bir örnek sunmuştur. Bu sayede, uzun vadede insan seviyesinde akıl yürütüebilen, kendi tecrübelerinden nedensel dersler çıkarabilen ve bunları anlaşılır şekilde açıklayabilen yapay zeka sistemlerine bir adım daha yaklaşmış bulunuyoruz.

Kaynaklar:

1. Hongjian Fang, Yi Zeng. "A Brain-Inspired Causal Reasoning Model Based on Spiking Neural Networks." *Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2021. Makalenin beyin-esinli SNN ile nedensel çıkarım çerçevesini tanımladığı orijinal çalışma.
2. Fang, H., & Zeng, Y. "A brain-inspired causal reasoning model based on spiking neural networks," *IJCNN 2021*. – BrainCog projesi kapsamında tanıtılan Causal Reasoning SNN (CRSNN) modülüne ilişkin açıklamalar ⁸ ²¹.
3. Fang, H., et al. "Brain-inspired Graph Spiking Neural Networks for Commonsense Knowledge Representation and Reasoning." *arXiv preprint arXiv:2207.05561*, 2022. – CRSNN modelinin genişletilerek ConceptNet gibi büyük ölçekli bilgi grafiklerine uygulandığı ileri çalışma; popülasyon kodlama, STDP ve ödül-modülatonlu STDP mekanizmalarını birleştirerek GNN benzeri bir SNN gerçekleştirdiklerini gösteriyor ¹⁶.
4. Frontiers in Neurorobotics, 2022 – "A brain-inspired robot pain model based on a spiking neural network" makalesi. CRSNN prensiplerinin robotik bir acı algısı senaryosuna uyarlanmasıı detaylandıran çalışma; STDP parametreleri ve popülasyon kodlama için Fang & Zeng (2021) referans alınmıştır ¹² ³⁰.
5. Sjöström, P. J., & Gerstner, W. (2010). "Spike-timing dependent plasticity." *Spike Timing-Dependent Plasticity*, **13**(1): 529-547. – STDP öğrenme kuralının parametrelerine ve biyofiziksel temellerine dair temel kaynak; modelde kullanılan STDP eğrilerinin şekli ve parametre seçimleri bu gibi literatüre dayanmaktadır ¹⁸.

[1 2 3 4 5 6 7 9 10 11 16 17 19 20 22 23 24 25 28 29 35 36 \[2207.05561\] Brain-inspired Graph Spiking Neural Networks for Commonsense Knowledge Representation and Reasoning](https://arxiv.org/pdf/2207.05561.pdf)
[https://arxiv.org/pdf/2207.05561](https://arxiv.org/pdf/2207.05561.pdf)

[8 13 14 15 21 \[2207.08533\] BrainCog: A Spiking Neural Network based Brain-inspired Cognitive Intelligence Engine for Brain-inspired AI and Brain Simulation](https://arxiv.org/pdf/2207.08533.pdf)
[https://arxiv.org/pdf/2207.08533](https://arxiv.org/pdf/2207.08533.pdf)

[12 18 30 31 32 33 Frontiers | A brain-inspired robot pain model based on a spiking neural network](https://www.frontiersin.org/journals/neurorobotics/articles/10.3389/fnbot.2022.1025338/full)
<https://www.frontiersin.org/journals/neurorobotics/articles/10.3389/fnbot.2022.1025338/full>

26 27 34 [2207.05561] Brain-inspired Graph Spiking Neural Networks for Commonsense Knowledge Representation and Reasoning
<https://arxiv.org/abs/2207.05561>