

sAraştırma Raporu

Veri Altyapısı ile Yapay Zekâ Modernizasyonunun Hızlandırılması

Sponsor: IBM Corporation

Ashish Nadkarni
Şubat 2021

Sriram Subramanian

Matt Leib

YÖNETİCİ ÖZETİ

Yapay zekâ (AI), makine öğrenmesi (ML) ve genellikle derin öğrenme (DL) yetenekleri artık dijital dönüşüm (DX) girişimlerinin temel bileşenleridir. Yapay zekâ ile elde edilebilecek iş fırsatları son derece umut vericidir. Rakiplerin müşteri tabanını büyütmek ve memnun etmek için daha önce mevcut olmayan çok sayıda içgörü ve yetenekler kazanmasıyla kuruluşlar, yapay zekâ temelinde hareket etmemenin potansiyel olarak bir iş felaketi olabileceğinin giderek daha fazla farkına varıyor. Bugün çok az sayıda işletme "yapay zekâ bize göre değil" ya da " yapay zekâ çoğunlukla aldatıcı bir şey" şeklinde bir inanışa sahip. Aksine çeşitli sektörlerde ve farklı boyutta şirketlerde dünyanın her yerinde ciddi yapay zekâ girişimleri görülüyor.

İşletmeler, deney aşamasından çıkıp yapay zekâ yatırımlarından iş değeri oluşturmaya kadar tüm süreci kapsayan, yapay zekânın yön verdiği bir dönüşüm ve modernizasyon girişiminin peşindedir. Yapay zekâyla ilgili dijital dönüşüm sürecinde iş yatırımlarının başarısı, uygun ölçekte yapay zekâ çözümleri geliştirmek, uygulamak ve sürdürmek için gereken uzmanlıkla doğrudan bağlantılıdır. Birçok kuruluşun iş kolu, BT çalışanları, veri uzmanları ve geliştiricileri yapay zekâ hakkında bilgi sahibi olmak, kullanım senaryolarını anlamak, işleri için bir yapay zekâ stratejisi belirlemek, ilk yapay zekâ girişimlerini başlatmak ve özellikle derin öğrenme olmak üzere makine öğrenmesi algoritmalarını kullanarak yeni içgörüler ve yetenekler sağlayan, elde ettikleri yapay zekâ uygulamalarını geliştirmek ve test etmek için çalışıyor.

Kuruluşların bu girişimleri ölçeklemesiyle yeni sorular ortaya çıkıyor. Standart, genel amaçlı hesaplamayı ve mevcut veya miras depolama altyapısını kullanamayacaklarını biliyorlar, aslında bunu bizzat deneyimlemiş de olabilirler. Ayrıca yapay zekâ eğitiminin (yapay zekâ modelinin eğitimi) ve çıkarımının (eğitilen modelin bir olayı anlama veya tahmin etme amacıyla kullanılması) eşit derecede ölçeklenebilir bir depolama altyapısıyla farklı türde ölçeklenebilir hesaplama gerektirdiğinin de bilincindedir.

IDC'ye göre altyapılarının bir kısmını veya tamamını modernize etmeden, mevcut altyapılarını kullanmaya çalışan kuruluşların hata payı daha yüksek. Modernizasyonda dahi odak noktası şirketten şirkete değişiyor. Hesaplama, bilgi işlem konusunda daha etkin olan işletmeler genellikle yapay zekâda depolamanın değerini azımsıyor. Bununla birlikte yapay zekâ uygulamaları ve özellikle olağanüstü büyük miktarda verileri ayrıştıran derin öğrenme sistemleri ciddi ölçüde zorlayıcıdır, çok sayıda hesaplama çekirdeğine dayalı, güçlü paralel işleme yetenekleri gerektirir ve standart depolama sistemleri bu yapay zekâ görevlerinin yürütülmesini yeterli

IDC'ye göre altyapılarının bir kısmını veya tamamını modernize etmeden, mevcut altyapılarını kullanmaya çalışan kuruluşların hata payı daha yüksek. Modernizasyonda dahi odak noktası şirketten şirkete değişiyor. Hesaplama, bilgi işlem konusunda daha etkin olan işletmeler genellikle yapay zekâda depolamanın değerini azımsıyor.

kadar sağlayamaz. Son olarak, yapay zekâyla ilgili girişimlerin Kubernetes ve/veya konteynerler ve hibrit bulut mimarisi üzerinden bir veya daha çok genel bulut hizmetiyle entegrasyonu içeren uygulama modernizasyonu çalışmalarını dikkate alması önemlidir.

IDC araştırması, depolama altyapısı bakımından, ayrıntılara doğru veya yeterli şekilde dikkat edilmemesinin yapay zekâ dönüşüm girişimlerini kısa sürede rayından çıkarabileceğini gösteriyor. Bu boşluğu kapatmak için, mevcut altyapıyla denemiş ve artık bunu üretime geçirmeye hazır olan işletmeler, gerekli paralel işleme performansını elde etmek için altyapılarını çok büyük ölçeklere çıkabilen ve hem küresel uygulama hem de veri erişimi için bulut, konteynerler ve performans yoğunluklu hesaplama entegre olan, daha modern depolama çözümlerine yatırım yaparak revize etmelidir. İşte bu noktada IBM Spectrum Scale ve IBM Elastic Storage System (ESS) bir yapay zekâ bilgi mimarisi için gereken bileşenleri sağlar. Özellikle yapay zekâ iş yüklerine odaklanan hibrit bulut devreye alımı, konteynerli devreye alma ve yapay zekâ iş yükleri için uygundur.

DURUMA GENEL BAKIŞ

Yapay Zekâ/Makine Öğrenmesi Bugün ve Yarın Bizimle

Dünyanın her yerinde işletmeler, dijital dönüşüm girişimlerine katalizör olarak yapay zekâ yatırımlarının sunduğu yeni fırsatları hevesle takip ediyor. Yapay zekâ soruları yanıtlamak, içgörülerini keşfetmek ve öneriler sunmak için doğal dilde işleme (NLP), resim/video analitiği, makine öğrenmesi, bilgi grafikleri ve diğer teknolojileri kullanan teknoloji grubudur. Bu sistemler, mevcut kanıta dayanarak olası yanıtları formüle eder, çok büyük hacimde içeriğin alınıp işlenmesiyle eğitilebilir ve yeniden eğitim ya da insan gözetimiyle hatalardan ve yanlışlardan ders çıkarıp adapte olabilir. IDC, 2022'ye kadar bu yapay zekâ merkezli kullanım senaryolarının en az %60'ının Global 2000 kuruluşunun en az %65'inde devreye alınacağını bekliyor, bu da 2019 yılına göre %34'lük bir büyümeyi gösteriyor.

IDC, 2022'ye kadar bu yapay zekâ merkezli kullanım senaryolarının en az %60'ının Global 2000 kuruluşunun en az %65'inde devreye alınacağını bekliyor, bu da 2019 yılına göre %34'lük bir büyümeyi gösteriyor.

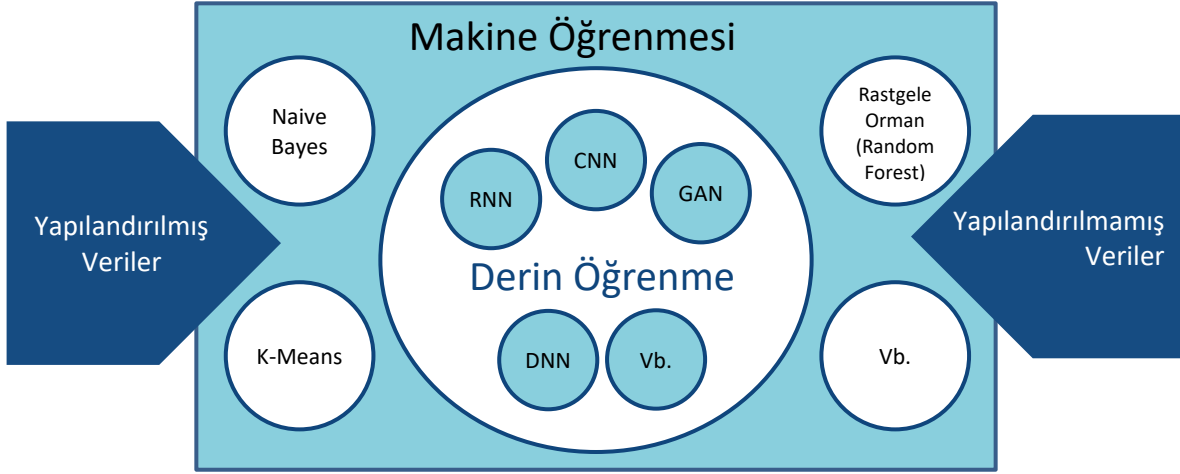
Yapay zekâ, süreçlerin ve iş akışlarının otomasyonunda kısa sürede işletmenin her alanında kendini gösteriyor. 2019'da IDC, müşteri deneyimi (CX), yasal ve kurumsal strateji, tesisler ve tedarik dâhil sekiz faaliyet alanında 176 dijital dönüşüm senaryosunu inceledi ve bu senaryoların yaklaşık %26'sının hem yapay zekâya bağlı olduğunu hem de halihazırda çoğu kuruluşta devreye alınmış olduğunu öngördü.

Bu da yakın bir zamanda önde gelen çoğu kuruluşun doğal dilde işleme, makine öğrenmesi, derin öğrenme ve konuşma-metin gibi yapay zekâ teknolojilerinden yararlanarak operasyonları ölçekleyeceği, yapılandırılmamış verileri anlayabileceği ve akıllı içgörüler sağlayabileceği anlamına geliyor. Bu arada yapay zekâ tabanlı kullanım senaryolarını kavram kanıtından (POC) üretim aşamasına nasıl taşıyabileceğini hâlâ anlayamayan kuruluşlar da geride kalacak ve aradaki dijital boşluk gitgide genişleyecek.

Makine öğrenmesi, bilgisayar sistemlerinin bir insan tarafından programlanmak zorunda kalmadan bir görev için davranışlarını öğrenmesini ve iyileştirmesini sağlayan yapay zekâ tekniklerinin alt grubudur. Makine öğrenmesi modelleri, bir görevi (örneğin bir insan yüzünü tanıma) "öğreninceye" kadar yapılandırılmış ve/veya yapılandırılmamış çok büyük hacimlerde verileri kullanarak kendilerini tekrar tekrar test ederek zaman içinde geliştirebilen algoritmalarıdır. Şekil 1'de derin öğrenmenin nasıl bir makine öğrenmesi alt grubu olduğu gösterilir. Tipik derin öğrenme mimarileri derin nöral ağlar (DNN), evrişimli sinirsel ağlar (CNN), yinelenen nöral ağlar (RNN), üretken çekişmeli ağlar (GAN) ve diğerleridir.

ŞEKİL 1

Makine Öğrenmesi ve Derin Öğrenme Uygulamaları



Kaynak: IDC, 2021

Yapay zekâ yazılım platformları dijital asistanlar gibi sohbete dayalı yapay zekâ yazılımları; verilerdeki gizli ilişkileri keşfetmek ve tahminlerde bulunmak için tahmine dayalı analitik; metni tanımak, anlamak ve değer çıkarmak için doğal dil ve metin analitiği; ses, görüntü ve konuşmadan bilgileri tanımak, belirlemek ve çıkarmak için ses/konuşma analitiği; model tanıma, nesnelere, renkler ve insanlar, yüzler, duygu, arabalar ve manzara gibi diğer özellikler dâhil resimlerden ve videodan tanıma, belirleme ve bilgi çıkarmaya yönelik resim ve görüntü analitiğini içerir.

Birçok işletme yapay zekâ girişimleriyle epey yol kat etmiş durumda ve üretim ölçeğinde yapay zekâyı devreye almaya başlamak için hazır oldukları bir aşamaya ulaştı. Diğerleri hâlâ yapay zekâyla ilgili deney aşamasında, bir üçüncü grup ise şu anda yapay zekâ uygulamalarının kuruluşları için ne anlama gelebileceğini değerlendirme sürecinde.

İlk grupta (devreye almaya hazır) ilgili olarak IDC, işletmelerin, devletlerin ve diğer kuruluşların uygulamaya başladıkları çeşitli yapay zekâ kullanım senaryolarını görüyor. Donanım, yazılım ve hizmetler bakımından işletmelerin harcadığı tutara göre sıralanan, bugünkü en yaygın beş kullanım senaryosu:

- **Otomatik müşteri hizmetleri temsilcileri.** Örneğin bankacılık sektöründe bu yapay zekâ uygulamaları, müşterinin ihtiyaçlarını ve sorunlarını anlayan ve bankanın müşteri sorunlarının çözümü için gereken kaynakları ve süreyi azaltmasına yardımcı olan bir öğrenme programı aracılığıyla müşteri hizmetleri sağlar. Bu temsilciler birçok sektörde yaygın olarak kullanılmaktadır.
- **Satış süreci önerileri ve otomasyon.** Çeşitli sektörlerde kullanılan bu uygulamalar, müşteri bağlamını gerçek zamanlı olarak anlamak ve satış temsilcilerine ilgili eylem önerileri sunmak için müşteri ilişkileri yönetimiyle (MİY) birlikte çalışan yapay zekâ uygulamalarıdır.
- **Otomatik tehdit istihbaratı ve önleme sistemleri.** Devletler ve sektörler genelinde tehditleri önlemenin önemli bir parçası olan bu yapay zekâ uygulamaları, istihbarat raporlarını işler, bunlardan bilgiler elde eder, farklı bilgi parçaları arasında ilişkiler kurar ve veritabanları, sistemler, web siteleri, vb için tehditleri tespit eder.

- **Dolandırıcılık analizi ve inceleme.** Diğer yerlerde yaygın kullanımının yanı sıra sigorta sektöründe bu yapay zekâ uygulamaları, dolandırıcılık işlemlerinin tespiti için kurallara dayalı öğrenmeyi kullanır ve sigortacılıkla ilgili çeşitli dolandırıcılık şemalarını belirlemeyi otomatik olarak öğrenir.
- **Otomatikleştirilmiş koruyucu bakım.** Üretim sektörlerinde bu yapay zekâ uygulamaları, kapalı kalma ve bakım maliyetini azaltan, olası fabrika ve makine arızalarının doğru tahmine dayalı modelini oluşturan makine öğrenmesi algoritmalarına dayalıdır.

Kuruluşlarda ilgi gören diğer yapay zekâ kullanım senaryoları (donanım, yazılım ve hizmetlerde yapılan harcama sırasında göre):

- Program danışmanları ve öneri sistemleri
- Teşhis ve tedavi sistemleri
- Akıllı işleme otomasyonu
- Kalite yönetimi incelemesi ve öneri sistemleri
- Kurumsal bilgi çalışanları için BT otomasyonu ve dijital asistanlar
- Uzman alışveriş danışmanları ve ürün önerileri
- Tedarik ve lojistik ve düzenleyici istihbarat
- Varlık/filo yönetimi ve otomatik talep işleme
- Dijital ikiz/gelişmiş Dijital simülasyon
- Kamu güvenliği ve acil durumda müdahale
- Uyarlanabilir öğrenme
- Akıllı ağ oluşturma
- Navlun, varlık ve filo yönetimi
- Farmasötik araştırma ve keşif

Yapay Zekâ Dönüşümü Güçlü Bir Veri Temeli Gerektirir

Makine Öğrenmesi ve Derin Öğrenme girişimleri ağırlıklı olarak çeşitli yapılandırılmış ve yapılandırılmamış veri girişleri birleşimine dayalıdır (yeniden Şekil 1'e bakın). Yapay zekâ projeleri veri toplama, yapılandırma, temizleme, doğrulama, model oluşturma, eğitime, test, çıkarım ve verilerin kullanımını sonlandırma dâhil birden çok adımı içerir. Daha hızlı erişim ve düşük gecikme süreli depolamadan düşük maliyet, arşiv depolamasına kadar her adımın değişen gereksinimlere sahip olmasıyla kuruluşların bu adımlar için uygun ve maliyeti karşılanabilir depolama altyapısını seçmesi gerekir. Ayrıca yaşam döngüsünde yapay zekâ veri kümesini yönetmek için farklı araçları da kullanmaları gerekir.

5G'nin kullanıma sunulması ve IoT tabanlı sensörlerin yaygınlaşmasıyla birlikte edge'de daha çok veri üretilir ve tüketilir. Daha hızlı yanıt ve daha üstün kullanıcı deneyimi sağlamak için cihazda, edge'de kameralar ve akıllı asistanlar devreye alınıyor. Endpoint'lere veya ağın ucunda işlenen yerel yapay zekâya olan ihtiyaç her geçen gün artıyor. Sınırlı bağlantırlığı olan edge cihazlarında çalışan, gecikme süresine duyarlı yapay zekâ uygulamaları yüksek düzeyde ölçeklenebilirlik gerektirir. Yapay zekâ destekli edge bilişim kullanım senaryoları benimsenir.

Bir Yapay Zekâ Veri Altyapısı Yapay Zekâ Dönüşümünü Hızlandırır

Kuruluşların yapay zekâyâ yatırım yapmasıyla birden çok altyapı tasarımı ve devreye alma faktörü de ciddi derece önemli hale geliyor. Hiper ölçek ve bulut hizmeti sağlayıcısı kılavuzundan bir şeyler öğrenerek kuruluşlar, birleşik bir veri altyapısı yaratılmasıyla altyapı gerekliliklerine yaklaşıyor. Bir veri altyapısı, son derece verimli ve ölçeklenebilir bilişim ve depolama katmanını içeren yapay zekâ girişimleri için ortak bir temeldir. Yapay zekâ iş yüklerine türdeş olarak yaklaşmak yerine bir veri altyapısı, bunları birleşik olarak ele alır ve bu iş yükünün bir kısmını yapılandırılmış ve yapılandırılmamış veri kümeleri karışımına dayalı uygun bir depolama katmanının desteklediği doğru hesaplama katmanına bağlar. Bu birleşik yaklaşım, üç boyutta yapay zekâ dönüşümüne ivme kazandırır:

- **Ölçek.** Ölçek boyutu, iş yükünün faaliyet gösterdiği ölçeği açıklar. Hesaplama, ağ ve veri sürekliliği (depolama) gibi temelin tüm alt boyutları donanımla ilgilidir. Orkestrasyon gibi yazılımla ilgili alt boyutlar, yığının boyutu ve karmaşıklığındaki artışla dengeyi korumak için eşitlilik elde ediyor.
- **Taşınabilirlik.** Bu çekirdek, edge ve endpoint uygulamalarında iş yükünün taşınabilmesi yeteneğidir. Günümüzde bu iş yüklerinin çoğu, doğası gereği statiktir (örneğin, tek bir devreye alımda çalışacak şekilde tasarlanmıştır). Şirketler her geçen gün bir devreye alımda iş yüklerini geliştirme (örn. genel bulut) ve bunları başka bir devreye alımda (örn. edge'de) kurmanın (üretimde) yollarını aramaktadır. Bu, şu anki mobil uygulama geliştirme ve devreye alma modeliyle kıyaslanabilir.
- **Zaman.** Bu boyut, iş yükünün kendisinin zaman sürekliliğiyle ilgilidir. Birçok yapay zekâ iş yükü, tasarımlarını yüksek performanslı bilişim veya büyük veri ve analitik devreye alımından alır; yapısı gereği toplu olacak şekilde tasarlanmıştır. Yapay zekâ iş yükleri her geçen gün ve yüksek performanslı hızlandırıcıların yaygınlaşması sayesinde akış verilerini gerçek zamanlı veya gerçek zamanlıya yakın şekilde analiz edebilir.

Yapay Zekâ Girişimleriyle İlgili Engeller

- IDC araştırması, bir veri yönetimi perspektifinden yapay zeka girişimleri konusunda işletmelerin genellikle aşağıdaki sorunlardan bahsettiğini gösteriyor:
- Çok uzun süren veri alma ve hazırlık döngüleri
 - Çeşitli analitik kullanım senaryoları için altyapı siloları
 - Tek bir gerçek kaynağı olmadan aynı verilerin birden çok kopyası
 - Yinelenebilirlik için veri menşesinin güvenli şekilde yönetimi ve korunması gerekliliği
 - Küresel erişilebilirlik (hibrit bulut) ve işbirliği ihtiyacı
 - Veriler yakalandığı ve düzenlendiği anda veri bütünlüğü

Yapay Zekâ Veri Altyapısının "Olmazsa Olmazları" ve Efsaneleri

Temel, buna karşılık genellikle göz ardı edilen temel depolama altyapısıdır. Yapay zekânın uygun ölçekte devreye alınması genellikle kapasite (büyüme) ve performans (IOPS ve bant genişliği) bakımından depolama altyapısına daha yüksek talepler getirir. Kuruluşlar diğer iş yükleri için kullanılan kurumsal depolamanın veya dâhili sunucu tabanlı depolamanın yapay zekâ uygulamalarını çalıştırmak için yeterli olduğunu varsayar. Ve altyapı oluşturulduğu anda depolamanın zincirin en zayıf halkası olduğunu farkeder. Bu yapay zekâ uygulamalarının her biri, farklı farklı gereklilikler doğurur ve dolayısıyla BT organizasyonu için çeşitli zorluklar yaratır. Bu nedenle BT satın alma uzmanları ve satıcı firmalar "Çekiç elindeyse her şey çivi olur" atasözünün açıkladığı durumdan kaçınmalıdır.

Veri altyapısına bütüncül olarak yaklaşmak daha tedbirli bir davranış olur. Veri sürekliliği ölçekleme ve erişim mekanizmaları minimum unsurlardır, diğer yandan kuruluşların hesaplama, depolama yazılımı ve sistem katmanları arasında entegrasyon ve ağ oluşturmayı dâhil edecek şekilde alanı genişletmesi

gerekir. İşletmelerin yalnızca diğer depolama sistemlerinin "takılmasını" içeren değil tutarlı, uçtan uca veri altyapısı mantığına geçmesi gerekir. IDC, özellikle depolamayla ilgili veri altyapısı gerekliliklerinin aşağıdaki bölümlerde üzerinde durulan temel alanlara kadar ayrıştırılabileceğine inanıyor.

Hesaplama Entegrasyonu

Tüm yapay zekâ iş yüklerinin genellikle konteynerli olduğu (kusurlu bir varsayım) varsayılr. Bunun dışında birçok yapay zekâ iş yükü, yalın donanımda çalıştırılır veya sanallaştırılmıştır. Özellikle yapay zekâ destekli uygulamalar genelde yalın donanımda çalıştırılır veya sanallaştırılmıştır. Çoğu yapay zekâ iş yükü, hızlandırıcıları kullanacak şekilde optimize edilir. Bu tüm yapay zekâ iş yüklerinin en iyi hızlandırılmış bilişimde çalıştığı anlamına gelmez, hızlandırılmış hesaplama genellikle iş yükleri için çeşitli zorlukları da beraberinde getirir.

Veri Sürekliliği ve Erişim

Yapay zekâ iş yükleri için hesaplama gereklilikleri değişirse veri sürekliliğine ilişkin gereklilikler de değişir. Yapay zekâ iş yükleri grubunun yeterince temsil edilmeyen ve yanlış anlaşılan bir yönü de veri sürekliliği katmanıdır. Bir diğer yanlış varsayım da tüm yapay zekâ iş yüklerinin büyük hacimde yüksek performanslı depolama gerektirdiğidir. Aslında tüm yapay zekâ iş yükleri "büyük veri kümeleri" değildir, kısa bir süre için çok sayıda küçük veri kümelerinin eş zamanlı örnekleri olabilir. Benzer şekilde, açık sistem bilişim platformlarında çalışan yalın donanım iş yükleri genellikle yatay ölçekleme bloğu veya dosya erişimini kullanır. Sanallaştırılmış iş yüklerinin hiper bağlantılı bir altyapıda (HCI) çalışması çok da nadir görülen bir durum değildir.

Çoklu protokol erişimi depolama altyapısına dâhil edilmiş, yapılandırılmış ve yapılandırılmamış veri karışımıyla birlikte sağlanır. Birçok Nesnelerin İnterneti (IoT) ve edge cihazı SMB veya NFS üzerinden iletişim kurar, birkaçı S3 kullanır. Akış verileri erişimi bazı durumlarda gereklidir. Bazı durumlarda yerel paralel dosya sistemi istemcisi de kullanılabilir.

Ölçekleme ve Katmanlama

Yapay zekâ ve makine öğrenmesi uygulamalarının desteklenmesi, depolama sistemlerinin uygun ölçekte performans sağlaması gerektiği anlamına gelir. Yapılandırılmamış veri havuzları için, ağ erişimine sahip paralel dosya sistemlerini kullanan depolama sistemleridir. Yapılandırılmış veriler için ise flash tabanlı depolama sistemlerinin kullanılmasıdır. Ölçekleme esas olarak yapay zekâ ve makine öğrenmesi uygulamalarının gerekliliklerini karşılamak için bir diğerinden bağımsız olarak performansı ve kapasiteyi artırmak veya azaltmaktır.

Ayrıca gelecekte olabilecek değişikliklerden etkilenmeyen altyapı için sistemin S3 gibi bilinen bir nesne depolama arayüzüyle düşük maliyetli bir nesne depolamasında daha eski veya daha soğuk olan verileri basit ve uygun maliyetli şekilde işleyebilmesi gerekir.

Yazılım Tanımlı Depolama

Yazılım tanımlı depolama için yapay zekâ ve makine öğrenmesi katalizör görevi görür. Donanımın üzerinde türdeş olmayan bir yazılım denetimi katmanı kod ve otomasyon olarak altyapıyı sağlar. Bu da yapay zekâ/makine öğrenmesi iş akışlarıyla entegrasyonu iyileştirmeye yardımcı olarak uygulamanın gereklilikleri doğrultusunda depolamanın sorunsuz olarak ölçeklenmesini sağlar.

Devreye Alma Çevikliği ve Esnekliği

Kullanım senaryolarını ele alan uygulamalar bir kuruluş tarafından özel olarak geliştirilebilir, ticari yapay zekâ yazılımına dayalı olabilir veya yapay zekâ SaaS olarak sağlanabilir. Özel geliştirilen ve ticari yazılım için devreye alma noktaları şirket içinde, IaaS platformunda bulutta veya hibrit bulut olmaktadır; diğer yandan şirket içi ortam ortak bir otomasyon ve orkestrasyon katmanı aracılığıyla genel bir bulut ortamıyla etkileşim kurar.

Yapay zekânın dağınık yapısı dikkate alındığında hesaplamaların verilerinin kaynağı veya üretildiği yere daha yakın bir konuma taşınmasının ideal olacağı varsayılabilir. Son zamanlarda çekirdek-edge-endpoint modeli yapay zekâyı açıklamanın genel geçer bir yöntemi (çekirdeğin bulutu içerdiği ve endpoint'lerin yerleşik istihbaratı içerdiği) haline geldi. Her konum için iş yükü profillerinin, dolayısıyla temel altyapı gerekliliklerinin değişiklik göstereceği unutulmamalıdır.

Çeşitli devreye alma senaryoları için çözümler şunları dikkate almalıdır:

- **Olağanüstü yüksek performanslı yapay zekâ modellerinin eğitimi için gereken veri hacminin güvenle işlenmesi.** Derin öğrenme eğitimi için performans gereklilikleri, yüksek bant genişlikli veri alımıyla birlikte GPU'ları kullanarak büyük ölçüde paralel işlemenin yürütülmesi yeteneğini kapsar.
- **Olağanüstü yüksek performansla yapay zekâ modelinin çıkarımı gerçekleştireceği veri hacminin güvenle işlenmesi.** Çıkarımla ilgili performans, eğitilen yapay zekâ modeliyle gelen verileri işleme ve gerçek zamanlıya yakın yapay zekâ içgörülerini veya kararlar sağlama yeteneği anlamına gelir.

Veri uzmanları ve geliştiriciler için, bulutta bir yapay zekâ girişiminin başlatılması bazen daha kolay olabilir, bu sayede derin öğrenme için genellikle hızlandırılması gereken, şirket içi hesaplamaların düzenlenmesi zorunluluğundan kurtulunmuş olur. Hızlandırılmış yapay zekâ bulut eşgörünümleri, genellikle açık kaynak yapay zekâ yazılım gruplarını içeren çoğu genel bulutta mevcuttur. Yapay zekâ eğitimi için hızlandırılmış bulut eşgörünümleriyle bulut hizmeti sağlayıcısı (SP) işlemciler, yardımcı işlemciler, ara bağlantılar, bellek boyutları, G/Ç bant genişliği vb. bakımından son kullanıcıya sunulacakları belirler. Tüm bulut hizmeti sağlayıcıları, bu bileşenlerin en optimize kombinasyonlarını sunmaz; bu da sonuç olarak veri uzmanlarının eğitim modellerini eğitebileceği kaliteyi ve hızı belirler. Sonuç olarak birçok kuruluş, şirket içinde devreye alımı seçer.

Son birkaç yılda yapay zekâ deneyleri sırasında birçok kuruluş, standart altyapıları veya temel bulut eşgörünümleriyle adeta "duvara topladı". Modellerin eğitimi çok uzun sürdü ve çıkarım süreci çok yavaştı. IDC araştırması, katılımcıların %77,1'inin şirket içi yapay zekâ altyapılarında bir veya daha çok sınırlamayla, %90,3'ünün ise bulutta hesaplama sınırlamalarıyla karşılaştıklarını söylediğini gösteriyor.

IBM STORAGE KULLANILABİLİR GENİŞLETİLEBİLİR KÜRESEL YAPAY ZEKÂ BİLGİ MİMARİSİ

Veri ve yapay zekâ için IBM Storage çözümleri, müşterilerin hibrit bulut ortamlarında üretim ölçeğinde yapay zekâ girişimlerini sorunsuz olarak tanıtmasını sağlar. IBM, yüksek performanslı yapay zekâ ve büyük veri çözümleri için verimli, güvenli, ölçeklenebilir ve yüksek kapasiteli depolamanın yanı sıra ölçeklenebilir yüksek performanslı iş yüklerinde liderliğini sürdürüyor. IBM'in depolama portföyü, edge'den temel veri merkezi ve genel buluta kadar entegre bir depolama ve veri yönetimi sağlayarak yapay zekâ modernizasyonunu hızlandırır. Kubernetes konteynerleri ve Red Hat OpenShift platformuyla entegrasyonla ve kapsamlı destekle birlikte gelir ve genel bulutta veya veri merkezi iş yükleri için devreye alınabilir ve erişilebilir. IBM Storage for data and AI, tüm kuruluş için uygun ölçekte

devreye alınabilen bir yapay zekâ bilgi mimarisiyle derin, artırılmış entegrasyon sağlayarak maliyeti ve karmaşıklığı azaltmayı amaçlar.

IBM Spectrum Scale

IBM Spectrum Scale, AI/ML, modelleme ve simülasyon ve analitik gibi performans yoğunluklu iş yükleri için tasarlanmış, dağıtık bir bilişim mimarisi temelinde dayanır. Basit yönetim ve ölçeklenebilirliğiyle işletim giderlerini, ilkeye dayalı veri optimizasyonu ve şeffaf veri yaşam döngüsü yönetimiyle sermaye giderlerini düşüren, kümelenmiş, paralel bir dosya sistemidir. IBM Spectrum Scale, tek bir dosya sistemine veya dosya sistemleri setine birden çok düğümden eşzamanlı erişim sağlar. Düğümler doğrudan bağlı, ağ bağlantılı, doğrudan ve ağ bağlantılı birlikte veya hiçbir şeyin ortak olmadığı bir küme yapılandırılmasında olabilir. Bu yatay ölçekleme çözümü ve üst düzeyde kullanılabilirlik platformu ortak veri kümelerine yüksek performanslı, paylaşılan erişim sağlar.

IBM Spectrum Scale veri eşlemesini, ilkeye dayalı depolama yönetimini ve birden çok yerde operasyonları destekler. BT operasyonları ekipleri Kubernetes konteyner düğümleri, IBM AIX düğümleri, IBM Z veya LinuxONE düğümleri, Linux düğümleri, Microsoft Windows Server düğümleri veya beşinin de karmasını içeren bir küme yaratabilir. IBM Spectrum Scale ortamlarda ortak paylaşılan veri erişimi sağlayarak sanallaştırılmış veya konteynerli eşgörünümlerde çalışabilir ve mantıksal bölümlere veya diğer hipervizörlerden yararlanabilir. Birden çok IBM Spectrum Scale kümesi, küresel veri işbirliği ve veri erişimi için geniş alan ağı (WAN) bağlantılarında veya bir konum içinde veri paylaşabilir. IBM Spectrum Scale, yüksek performanslı bilişim sektöründe uzun yıllar verdiği hizmeti temel alarak, yapay zekâ altyapısı için son derece güçlü bir temel sağlar (bkz. Şekil 2).

IBM Spectrum Scale, yüksek performanslı bilişim sektöründe uzun yıllar verdiği hizmeti temel alarak, yapay zekâ altyapısı için son derece güçlü bir temel sağlar.

ŞEKİL 2

IBM Spectrum Scale ve Hibrit Bulut Veri Hizmetleri



Kaynak: IDC, 2021

IBM Spectrum Scale küresel bir ad alanı, IBM Spectrum Scale kümeleri arasında paylaşılan dosya sistemi, birden çok düğümden eşzamanlı dosya erişimi, replikasyonla üst düzeyde kurtarılabirlik ve veri erişimi, bir dosya sistemi bağlandığında değışiklik yapma yeteneđi ve büyük ortamlarda dahi basitleştirilmiş yönetim sağlar. IBM Spectrum Scale'in başlıca farklılaştırıcı özellikleri:

- Bir WAN genelinde veya bir konumda ayrı kümeler arasında veri paylaşımına izin veren IBM Spectrum Scale kümelerinde paylaşılan dosya sistemi erişimi
- Sistem performansını iyileştiren, patentli bir paralel dosya sistemi kullanarak iyileştirilmiş sistem performansı
- Belirteç yönetimi kullanarak istemcilere eşzamanlı ve ayrıntılı erişim aracılığıyla dosya tutarlılığı
- Dosya sistemi denetim günlüğü gibi özelliklerle artan veri erişilebilirliği ve güvenilirliği ve uzun mesafeye yayılabilen akıllı bağlantılar gibi yapılandırılabilir özellikler
- Dosya sistemi bağlı olduğunda disk veya sunucu kaynaklarının eklenmesini veya silinmesini sağlayan gelişmiş sistem esnekliği
- Flaştan HDD'lere buluttan manyetik banta ve hatta ilkeye dayalı veri azaltımı dâhil, güçlü ilke odaklı, otomatikleştirilmiş katmanlı depolama yönetimiyle bilgi yaşam döngüsü yönetimi (ILM) sağlamaya yardımcı olacak basitleştirilmiş depolama yönetimi
- Doğrudan çoğu uygulamadan yürütülebilen birçok standart dosya sistemi arayüzüyle basitleştirilmiş yönetim
- Veri erişilebilirliği, bütünlük, güvenlik ve optimize edilmiş konteyner yapıllı depolama ve Red Hat OpenShift ile entegrasyon sağlayan hibrit bulut devreye alımı.

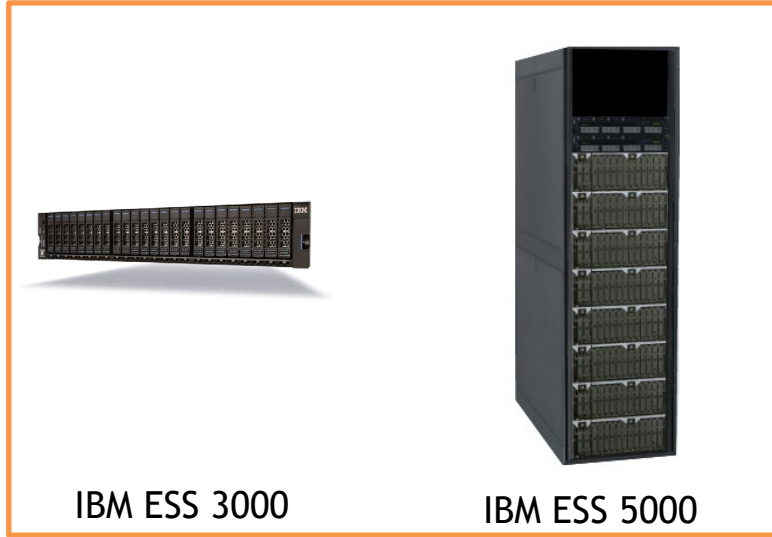
IBM Elastic Storage System

IBM Elastic Storage System, yapı bloklarının daha kolay yapılandırması ve yönetimini sağlayacak şekilde tasarlanmış, yazılım tanımlı depolamanın modern bir uygulaması olup BT kuruluşlarının yapay zekâ, büyük veri ve analitik uygulamaları dâhil performans yoğunluklu bilişim uygulamaları için hızlı, üst düzeyde ölçeklenebilir Depolamayı devreye almasını kolaylaştırır (bkz. Şekil 3). IBM ESS:

- Eksabayt düzeyinde ölçeklenebilirlik ve tüm altyapıda tutarlı hizmet kalitesi sağlamak için NVMe flash depolamasıyla oluşturuldu
- Birleşik bir küresel depolama sistemi sağlamak için IBM Spectrum Scale'ın dosya yönetimi ve veri hizmetleri yetenekleriyle entegre edilebilir
- İşletmelerin uçtan çekirdek verilerine kadar depolama gerekliliklerini birleştirmesini ve genel bulutla entegrasyonunu sağlar, bu sayede verimsizliği azaltır, edinme maliyetlerini düşürür, depolama yönetimini basitleştirir ve veri silolarını ortadan kaldırır
- Yalın donanım olarak veya sanallaştırılmış ortamlarda devreye alınan birden çok zorlu uygulama için tutarlı, yüksek performanslı erişim sağlar (Kubernetes ve Red Hat OpenShift entegrasyonunu destekleyerek bulut tabanlı uygulamaların devreye alınmasını kolaylaştırır.)

ŞEKİL 3

IBM Elastic Storage System



Kaynak: IDC, 2021

IBM ESS, 3000 ve 5000 olarak iki biçim katsayısı olarak mevcuttur:

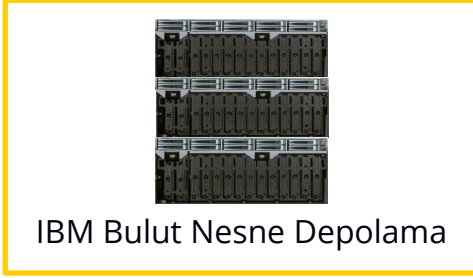
- IBM Elastic Storage System 3000 (ESS 3000), analitik için veri yönetimi zorluğunun karşılanması ve aşılması amacıyla tasarlandı. Kompakt 2U kasada paketlenmiş olan ESS 3000, tümü NVMe depolama ve basit, hızlı konteynerli yazılım kurulumu ve yükseltme sayesinde yapay zekâ/derin öğrenme ve performans yoğunluklu uygulamalar için değer elde etme sürecini hızlandırır. ESS 3000'in donanım ve yazılım tasarımı, veri açıklığı olan bilişimin tam anlamıyla kullanımını sağlamak için gereken sektörde lider performansa kurumsal erişimi sağlar.
- IBM ESS 5000 yazılım tanımlı IBM Spectrum Scale depolamasını IBM POWER9 işlemci tabanlı G/Ç yoğunluklu sunucularla bir araya getirerek yüksek verimle yatay ölçekleme petabayt kapasiteli düğümler sunar. Kuruluş genelinde depolama gerekliliklerini IBM ESS 5000 ve NVMe tabanlı ESS 3000 ile birleştirerek BT ekipleri verimliliği azaltabilir, edinme maliyetlerini düşürebilir, sağlık, medya, kamu ve finans hizmetleri alanında tipik olarak görülen, zorlu yapay zekâ, HPC, analitik ve/veya yüksek kapasiteli depolama gerekliliklerini destekleyebilir. ESS 5000, terabaytlarla başlayıp yüzlerce petabayta kadar büyüyebilir veya yüksek maliyetli veri silolarını ortadan kaldıran tek bir ad alanıyla eksabayt ölçeğine çıkabilir. IBM Spectrum Scale, büyüdükçe verimi artıran IBM ESS 5000'in merkezindeki paralel dosya sistemidir. Yatırım koruması için önceki ESS modelleriyle entegrasyonu sağlar ve bulut depolama ve IBM Tape gibi düşük maliyetli seçenekler sunar. IBM Spectrum Scale sayesinde BT, veri silolarını ve sorunlu noktaları ortadan kaldırabilir, depolama yönetimini basitleştirebilir ve verilere daha hızlı erişim elde edebilir.

IBM Cloud Object Storage

IBM Cloud Object Storage (COS) yapılandırılmamış verilerin edge (uç), çekirdek veri merkezi ve özel veya genel bulutlarda depolanmasına yönelik, sektörde önde gelen, yazılım tanımlı, üst düzeyde ölçeklenebilir ve uygun maliyetli bir Depolama çözümdür. IBM Cloud Object Storage, yapay zekâ, analitik, IoT, video ve görüntü havuzları için performans yoğunluklu altyapının devreye alınması veya

modernizasyonu için idealdir. Verimi 12 düğümlü bir kümede 55 GB/s seviyesine kadar artırırken yeni 18 TB kaplamalı manyetik rezonans (SMR) kullanarak kuruluşların depolama maliyetini %12'ye kadar düşürmesini sağlayan, benzeri görülmemiş bir değer de sağlar. İşletmeler, zorlu veri gölleri ve büyük kapasite gereklilikleri için özelleştirilebilir, IBM'in yerel veya değişik bölgelere dağıtılmış veri korumasıyla verilerini koruyabilir (bkz. Şekil 4).

ŞEKİL 4



Kaynak: IDC, 2021

IBM Cloud Object Storage aşağıdaki gibi önemli avantajlarla yapay zekâ altyapısı için temel bir sistemdir:

- **Ölçeklenebilirlik:** COS verilerde olağanüstü büyümeyi destekleyerek performansı ve kapasiteyi terabaytlardan eksabaytlara kadar ölçekler.
- **Güvenlik:** COS yerleşik şifreleme ve ilkeye dayalı kilitlenebilir WORM (bir kere yazılır bellek) depolamasına sahiptir.
- **Basitlik:** COS her konumdan verilere eşzamanlı olarak erişebilir ve otomatik olarak yedek sisteme geçiş, verileri yeniden oluşturma, otomatik genişletme ve yeniden dengeleme sağlar.
- **Tasarruf:** COS, Information Dispersal Algorithm (IDA) verimliliğiyle coğrafi olarak korunan veriler sağlar ve yalnızca yazılım veya tamamen desteklenen aygıt çözümleri olarak kullanılabilir.
- **Arama yetenekleri:** COS zaman tasarrufu için özel içgörüler ve arama olanağı sunar. Kuruluşlar, değeri artırmak için özel meta veriler yaratabilir.
- **Gelişmiş dosya erişimi:** Dosya tabanlı uygulamaları nesne depolamasına kolaylıkla bağlayarak yeni bir dosya erişim yazılımı ağı geçidi, SMB veya NFS erişimi kullanarak herhangi bir Windows veya Linux dosya sistemine sorunsuz şekilde eklenebilir.
- **Yüksek hızlı aktarım:** IBM Aspera yüksek hızlı veri aktarım seçeneği verilerin aktarımını kolaylaştırır ve esnek depolama sınıfı katmanları veri erişimi gerekliliklerini karşılarken maliyetlerin yönetimine yardımcı olur.

GELECEĞE BAKIŞ

2022'de küresel GSYİH'nin toplam %65'i dijitalleşmiş olacak ve 2020-2023 için dünya genelinde BT harcaması 6,8 milyar \$ olacak (bkz. *IDC FutureScape: Worldwide Digital Transformation 2021 Predictions*, IDC #US46880818, Ekim 2020). Dijital altyapı, geleneksel merkezi kurumsal hizmetlerle veya bağımsız bulut veri merkezleriyle sınırlı değildir. Uygulamaların değişimini ve dönüşümün şifresini sağlayan tüm kaynakları ve varlıkları içerir. Müşteri deneyiminin iyileştirme temeli olacaktır. İş

operasyonlarına istihbaratın/otomasyonun dâhil edilmesini de sağlar ve kuruluşun ve sektörün dijital tarafı için doğrudan sürekli inovasyonu destekler. Başarılı bir dijital strateji siloları ortadan kaldırmak, teknolojinin getirdiği engelleri aşmak ve yalnızca geleneksel araçları ve uygulamaları desteklemenin ötesine geçmek için dijital altyapıyı dönüştürmelidir.

IDC, yapay zekânın dijital altyapının temeli olacağına inanıyor. IDC'nin Müşteri İlgileri ve Analiz Grubu son zamanlarda çeşitli sektörlerden her boyutta kuruluştaki şu an ve gelecek için BT harcamalarını ve benimseme planlarını anlamak amacıyla bir araştırma gerçekleştirdi. Araştırmaya göre dünya genelinde çeşitli sektörlerden BT kuruluşlarında 3600 BT karar alıcısının yaklaşık %76'sı yapay zekânın DX stratejilerinin bir parçası olduğunu veya 1-2 yıl içinde bu yönde bir beklentisinin olduğunu belirtti. Araştırmaya katılanların yalnızca %22'si yapay zekânın önümüzdeki 3-5 yıl içinde DX stratejilerinin önemli bir parçası olacağını ifade etti. Önümüzdeki 1-2 yıl içinde yapay zekânın DX stratejilerinin temel bir bileşeni olmasını bekleyenler arasında telekomünikasyon, doğal kaynaklar, eğitim ve profesyonel hizmet kuruluşları dijital dönüşüm çalışmalarında yapay zekâyâ yönelme olasılığı en yüksek olanlar.

Yapay zekâ, şirketlerin mümkün olan en kısa sürede verilerinden elde edebileceği "değer elde etme süresiyle" alakalı. IDC'nin yapay zekâ konulu FutureScape raporunda "yapay zekânın çağımızın en köklü inovasyonu" olduğu öne sürüldü. Yapay zekâ artık sadece "olsa iyi olur" dediğimiz bir şey değil. Dünyanın içinde bulunduğu pandemi yapay zekânın benimsenme sürecini hızlandırdı ve yapay zekâ tüm iş süreçlerinde görünür hale geldi. Makine öğrenmesi, sohbete dayalı yapay zekâ ve bilgisayar görüşü destekli yapay zekâ çözümleri işte dirençlilik, hızlandırılmış inovasyon ve dönüşüm niteliğinde müşteri ve çalışan deneyimlerinin ön planında yer alıyor. Bahse konu araştırmaya katılanların yaklaşık %51'i, şu anda yapay zekâyı değerlendirdiklerini veya halihazırda üretimde olduğunu belirtti, bu oran 2019'daki %34'e göre yükselmiş. Yapay zekâ en çok çalışanların işlerinde daha başarılı olmasına yardımcı olma noktasında etkili. Tam devreye alımlardan elde edilen avantajların daha somut hale gelmesiyle yapay zekânın kuruluşlardaki artışı devam edecek.

IDC, yapay zekâ altyapısı yatırımlarının önümüzdeki birkaç yılda güçlenmeye devam edeceğini tahmin ediyor. IDC'nin öngörülerine göre yapay zekâ donanımı (sunucu ve depolama birlikte) gelirleri %10,3'lük yıllık artışla 2020'de 13,4 milyar \$ değerine ulaşacak. Donanım pazarında, yapay zekâ depolamasının 2020'de %11,4 büyümesi bekleniyor. Genel donanım pazarının 2021'de %35,5'lik yıllık büyümeyle güçlü bir geri dönüş yaşaması, yapay zekâ depolamasının yıllık %43,1 büyümesi bekleniyor.

Bu depolamanın önemli bir kısmı iş yükü katmanının çekirdekte buluta veya uca ve ardından aksi yönde taşınmasını sağlayan sorunsuz bir mobilite katmanına sahip hibrit bulut ortamlarında devreye alınacak. Ortak bir erişim ve denetim düzlemi sağlarken verilere doğru hesaplama hareketini desteklediğinden hibrit bulut için depolama önemli bir temel olacak.

Depolama, ve özellikle hibrit bulut ortamlarındaki depolama, yapay zekâ girişimlerinin bugün ve gelecekte ölçeklenebileceği temel prensip olmaya devam edecek. Yapay zekâ yatırımları ve yapay zekâ nedeniyle veri modernizasyonu girişimleri, yatay ölçeklenebilir dosya depolama ve yapılandırılmamış verilere yatırımları destekleyecek. IDC son günlerde BT altyapısı benimseme trendlerini belirlemek için dünya genelinde 624 BT uygulayıcısı ve operasyon ekibi arasında bir araştırma yaptı. Bu araştırmada IDC, katılımcıların %65'inden fazlasının yapay zekâ gibi yüksek performanslı iş yükleri için NFS aracılığıyla veya yerel

Yapay zekâ depolamasının 2020'de %11,4 büyüyeceği tahmin ediliyor. Genel donanım pazarının 2021'de %35,5'lik yıllık büyümeyle güçlü bir geri dönüş yaşaması, yapay zekâ depolamasının yıllık %43,1 büyümesi bekleniyor.

olarak erişilen yatay ölçeklenebilir dosya sistemlerini tercih ettiğini gördü. Aslında, özellikle eğitim ve çıkarım iş yüklerini içeren yapay zekâ iş yükleri için performans, en önemli depolama gerekliliğiydi. Bunun ardından hibrit bulutta devreye alma kolaylığı ve hizmet kalitesi geliyordu.

BT SATIN ALMA UZMANLARI İÇİN TEMEL KILAVUZ

Teknoloji satın alma uzmanlarının haklı olarak kendi yapay zekâ altyapı yazılım grubunu oluşturma süreci hakkında kafası karışık. Kullanım senaryolarını belirlediler, yapay zekâ girişimlerini başlattılar ve veri uzmanları ve uygulama geliştiricilerini işe aldılar veya eğittiler, ancak bir anda kendilerini yapay zekâ modellerinin geliştirileceği altyapı konusunda kısıtlanmış olarak buluyorlar. Genellikle mevcut altyapıdan kısa bir süre yararlanılır, bunun ardından hızlandırılmış altyapı yatırımları gelir. Veri uzmanları, kendilerini hızlandırılmış altyapıda yığınları bir araya getirirken ve bunu çalıştırmaya uğraşırken buluyor, oysa bu iş tanımlarında yer alan bir bölüm değil. BT altyapısı ekipleri, veri uzmanlarının ihtiyacı olan yazılım gruplarına aşına değil ve bunları bir araya getirip optimize edemiyor. Bunun sonucunda sunucu tedarikçileri ve bulut hizmeti sağlayıcılarının kendi bildikleri şekilde, kendi yazılım gruplarıyla doldurmaya çalıştıkları ciddi bir beceri açıklığı oluşuyor. Günümüzde aynı değer zincirinde birden çok üye tarafından geliştirildiğinde birbiriyle örtüşen, satıcı firma kadar yazılım grubu vardır.

İş Sonuçlarıyla Başlayın

Kuruluşlar, yapay zekâ altyapısı yatırımından yararlanan iş sonuçlarını belirlemek için kullanım senaryolarını ve hizmet kısıtlamalarını bir araya getirerek işe başlamalıdır. Bu tür yatırımların avantajlarını ölçmeleri ve niceliğini değerlendirmeleri gerekir. Örneğin, rekabette farklılık arayışındayken ne kadar ve ne zamana kadar soruları sorulmalıdır. Bu ölçütler daha sonra bir uygulama mimarisi seçimine yol açmalıdır. İşletmeler müşteri hassasiyetini, isteklerini ve ihtiyaçlarını daha iyi anlayarak ve bu konuda duyarlı olmaya odaklanarak markalarını geliştirmenin yollarını ararken yapay zekâyı düşünmelidir, bunun sonucunda da artan gelir ve kârlılık avantajı doğar.

Bütüncül Bir Yaklaşımı Benimseyin

Bir yapay zekâ girişimini uygularken resmin bütününe (kapsamlı bir görünüm) bakmak önemlidir. Tek başına yalnızca bir soruna bakmak, başka bir silonun daha yaratılmasıyla veya daha da kötüsü birden çok mimari ve çözüm arasında entegrasyon ve birlikte çalışabilirlik özelliğinin eksikliği nedeniyle ortamdaki karmaşıklığın artmasıyla sonuçlanabilir. Bir veri altyapısına uçtan çekerdeğe ve buluta kadar ve yapay zekâ ve diğer kritik misyonlu uygulamalar gibi çeşitli kullanım senaryolarında küresel bir çözüm olarak bakılmalıdır. Çekerdekte, uçta (edge) ve bulutta yapay zekâ ve diğer kritik misyonlu uygulamalar gibi kullanım senaryolarına hizmet eden, genel bir çözüm işlevi görmelidir.

Doğru Uygulama ve Veri Mimarisini Geliştirin

Bir yapay zekâ uygulaması ve veri mimarisi geliştirmek karmaşık bir görevdir. İsteki gerekliliklerin ve sonuçların kararlı, yapay zekâ destekli bir iş akışına dönüştürülmesi sürecini kapsar. İş akışı, yapay zekâ yeteneklerinin uygulama davranışını nasıl geliştirdiğini, verilerin nasıl alındığını ve analiz edildiğini ve uygulamanın diğer iş uygulamaları ve kullanıcılarla nasıl etkileşim kurduğunu açıklamalıdır. Burada verilerin uygulamalar tarafından nasıl kullanıldığı, üretildiği, analiz edildiği ve donanımdaki etkilerine odaklanılmalıdır. Sıfırdan bir plan oluştururken özel (açık kaynak veya patentli) ve hazır yazılım bileşenleri arasındaki karışıma önem verilmelidir.

Doğru Referans Grubu Seçin

Bazı satıcılar ve hizmet sağlayıcıları, bir yapay zekâ altyapısının uygulanması için referans grupları öne sürdü. Birçoğu yapısı gereği "açık" olduğundan modüler bir "tak-çalıştır" deneyimi sağlıyor ve sermaye gideri bakımından olumlu şekilde "kullandığın kadar öde" esasına dayalı bir hizmetle kullanılabilir. Yapay zekâ altyapı yatırımlarının kısa sürede yüksek maliyetli olabileceği düşünüldüğünde bu önemli bir nokta. IDC, çıkaracağı bir raporda sık kullanılan tedarikçi referans gruplarıyla ilgili perspektifini açıklamayı planlıyor.

Referans grupları incelerken BT açısından dikkate alınacak avantajlar şunlar: azalan maliyetler, veri ve uygulama kullanılabilirliği, etkin altyapı kullanımı ve birleştirmesi ve, mümkünse, birlikte çalışılabilirlik özelliğine sahip tek bir uygulama teslim platformu.

Yapay Zekâ İçin Bir Bilgi Mimarisi Oluşturun

Kuruluşların bulunduğu yerden ve referans mimariyle ilgili sorunları ele aldığı yerden bağımsız olarak, her türden tüm verilere esnek, organize erişim sağlayacak bir veri yönetimi stratejisine ihtiyacı vardır. Bir modernizasyon girişimi, diğer bulut platformlarıyla iletişim kurabilen, esneklik ve seçim olanağı olan, açık, genişletilebilir bir temel sağlayan bir bilgi mimarisi tanımlar ve devreye alır. IBM'in yapay zekâ yolculuğunu hızlandırmaya yönelik hibrit veri yönetimi stratejisi, toplama, düzenleme, analiz ve eklemeyen oluşan dört adımlı bir yapay zekâ merdiveniyle tanımlanan normatif bir yaklaşımdır.

- **Toplama:** Herhangi bir veritabanı veya depolama tesisinden doğru yerde verileri basit ve erişilebilir hale getirin.
- **Düzenleme:** Verilerin bilgi yaşam döngüsünün tüm aşamalarında güvenilir, eksiksiz ve tutarlı olmasını sağlayın: veri profili oluşturun, temizleyin ve kataloğlara ayırın; koruma ve uyumu sağlayın; ilkeye dayalı görünürlük, tespit ve raporlamayı sağlayın.
- **Analiz:** Hem yapılandırılmış hem yapılandırılmamış verileri keşfetmek ve analiz etmek ve bunları güvenle devreye almak için entegre araçları kullanarak yapay zekâ modelleri oluşturun, devreye alın ve yönetin.
- **Ekleme:** Sağlanan çözümleri ve hizmetleri kullanarak modelin önerdiği kararlarda güveni ve şeffaflığı sağlayın, kararları açıklayın, tarafsızlığı tespit edin.

Yapay zekâ yolculuğu, kuruluşun tümüne kolaylıkla yayılabilen bir bilgi mimarisiyle verilerin özümsemesinden içgörülere doğru taşınmasıyla ilgilidir. Yapay zekâ merdiveninin her parçasının tüm yolculuğa bir entegrasyon sağlaması önemlidir. Depolama genellikle birlikte veya kapsamlı bir altyapı çözümleri grubuyla entegre edilmeyen çözümler ve veri siloları yaratan bazı depolama çözümleriyle ustalıkla uygulanmıştır. Müşteriler, büyük bir dosya veya nesne depolama sisteminde verileri depolayabilir, ancak bu durumda bu verilerle ilgili ayrıntılara sahip olmayabilir veya daha fazla içgörü için bu verileri kullanmayabilir. Müşteriler yine başlayabilir veya yolculuğun bir kısmına bir projeyi odaklayabilir, ancak her proje kaynakları optimize etmek için genel bir yapay zekâ bilgi mimarisini düşünmeli ve genişleyen yapay zekâ iş yükleri için altyapınızı modernize etmelidir.

Yapay zekâya yatay ölçeklenebilir dosya sistemleri, türdeş olmayan bilişim ve dağınık hesaplama ve depolama erişimi için yüksek hızlı ara bağlantılar gibi önemli öğeleri ödünç alan, yüksek performanslı bileşik bir uygulama (birbirine bağlı birkaç uygulamadan yapılmış) gibi davranan kuruluşlar, sonuç olarak yapay zekâ altyapılarını ölçekleyebilenlerdir.

Doğru Ortaklıklardan Yararlanın

BT satın alma uzmanları için, uçtan uca çözüm sağlayıcısıyla ortaklık kurmak uzun vadede başarı için önemlidir; ancak IDC şu anda pazardaki herhangi bir tedarikçinin henüz bu uçtan uca ortamı sağladığına inanmıyor, tedarikçiler bu aşamaya gelmek için çok çalışsa da... Buna karşılık güvenli bir ortakla birlikte kuruluşlar, işlerini büyütmek için yapay zekâ yaklaşımlarını daha iyi kullanabilir. Çevik ve uyum sağlayabilir hale gelebilir ve karlılık elde etmek için içerideki sinerjilerden yararlanabilir. Bunun sonucunda kendilerini yeniden keşfederek sektördeki yıkıcı güçleri geride bırakabilir. İdeal bir ortak şunları sağlar:

- Küçük çaplı laboratuvar uygulamalarından büyük ölçekli devreye alımlara kadar ölçeklenen, kanıtlanmış çözümler
- İş odağına uygun segmentlere önem veren, dikey segment uzmanlığı
- ISV ve altyapı sağlayıcılarından oluşan çeşitli bir ekosisteme entegre erişim
- Yeni veri kaynaklarına yatırımlarının uzun vadeli değerini korumanızı ve en üst düzeye çıkarmanızı sağlayacak, veri öncelikli bir bakış açısı
- Büyük projelerin yazılım, donanım ve güvenlik özelliklerini basitleştirme konusunda doğrulanan başarı

ZORLUKLAR/FIRSATLAR

Kuruluşlar İçin

Bu teknik belgede üretim için yapay zekâ uygulamalarını ölçeklemeye hazır olduğunda kuruluşların karşılaştığı çeşitli zorluklar ele alınır. Veri hazırlığından model geliştirmeye ve çalışma zamanı ortamlarına yapay zekâ modellerinin eğitime, devreye alınmasına ve yönetimine kadar temel altyapının gereklilikleri, genel amaçlı donanımın eski modellerine meydan okur. Üstün performans, ölçekleme, veri erişimi ve entegrasyonla veri yoğunluklu iş yükleri için tasarlanmış ve hibrit bir bulut ortamıyla karıştırılabilen bir altyapıya yapılan yatırımlar uzun vadeli değer ve hizmet kalitesi sağlar. Kuruluşların mevcut genel amaçlı depolama platformlarını yapay zekâya özgü işleme görevlerine uygun depolama sistemleriyle değiştirme veya bunlarla tamamlama yönünde karar alması gerekir. Son teknoloji yapay zekâ uygulamalarının geliştirilmesi ve çalıştırılması temelini bu oluşturacak.

IBM İçin

IBM'in için zorluğu her zaman pazarda bilinen bir konudur. IBM, altyapı yazılım gruplarıyla (Red Hat OpenShift gibi) ve genel bulutla entegre olan, ancak potansiyel müşterilerin yanılıp karmaşı veya daha yüksek maliyetli olarak algıladığı, olağanüstü ve kapsamlı yapay zekâ altyapı çözümleri sunar. Olağanüstü veri yoğunluğu olan iş yükleri için büyük ticari depolama sistemi tedarikçilerinden birini seçmeye karşılık hemen ardından düşünmeden verilecek tepki de bu kuruluşların gerçekten yarar sağlayabilecekleri yapay zekâ altyapı çözümlerinden mahrum bırakılması olur. Örneğin IBM Spectrum Scale ve IBM Elastic Storage System ile IBM Storage for data and AI solutions, en büyük veri merkezlerinin birçoğu için bir süperbilgisayar yapı bloğudur. Yeni yapay zekâ iş yüklerinin HPC gereklilikleriyle süper bilişim uygulamalarını ciddi ölçüde taklit etmeye ve üzerinde çalıştıkları bulut altyapısı ve şirket içi altyapıya meydan okumasıyla birlikte IBM'in sahnedeki yerini alıp yeni müşteriler kazanma sırası gelmiştir.

SONUÇ

Son birkaç yılda IDC, çeşitli yapay zekâ yeteneklerini geliştirmeye nasıl başladıkları da dâhil birçok kuruluştta yapay zekâ dönüşümüne şahit oldu. Nispeten deneyimsiz çalışanlar tarafından başlangıçta deney olarak başlatılan ve mevcut olan herhangi bir altyapıda gerçekleştirilen bu girişimler artık yayılmaya başladı. Birçok kuruluş kapsamlı yapay zekâ uzmanlığı geliştirdi ve yapay zekâ yeteneklerinin işlerinin önemli bir parçası haline gelme hızını bizzat deneyimliyor.

Diğer yandan yapay zekânın çalıştırılacağı altyapı konusunda bir öğrenme eğrisiyle BT de bir dönüşüm sürecinden geçti. Günümüzde derin öğrenme eğitimi ve çıkarıma ilişkin altyapı gereklilikleri ve bu ortamların üretim için nasıl ölçekleneceğine ilişkin çok daha fazla netlik var. Derin öğrenme eğitiminin diğer uygulamalardan farklı bir altyapı gerektirdiği oldukça bariz. Derin öğrenme eğitiminde güçlü işlemciler, güçlü yardımcı işlemciler, ölçeklenebilir depolama destekli hızlı ara bağlantılar, büyük G/Ç bant genişliği ve çok fazla bellek kapasitesine sahip kümelenmiş düğümler istenir.

Bugün, BT'nin alması gereken en büyük karar; birinci sınıf sistemlerle veri altyapısı yapay zekâ uygulamalarını en iyi şekilde nasıl tasarlayıp devreye alabileceği ve bunları nasıl optimize edebileceği yönünde olmalıdır.

IDC; Spectrum Scale, Elastic Storage System ve Cloud Object Storage'ın aralarında olduğu IBM veri ve yapay zekâ için depolama çözümlerinin benzersiz bir değer ve performans sunduğuna inanıyor.

Bugün, BT'nin alması gereken en büyük karar; birinci sınıf sistemlerle veri altyapısı yapay zekâ uygulamalarını en iyi şekilde nasıl tasarlayıp devreye alabileceği ve bunları nasıl optimize edebileceği yönünde olmalıdır.

IDC Hakkında

International Data Corporation (IDC) bilgi teknolojisi, telekomünikasyon ve tüketici teknolojisi pazarlarında lider küresel bir pazar istihbaratı, danışmanlık hizmetleri ve etkinlikleri sağlayıcısıdır. IDC, teknoloji alımları ve iş stratejisi konusunda gerçeklere dayalı kararlar alınmasında BT uzmanlarına, şirket yöneticilerine ve yatırım topluluğuna yardımcı olur. Sayısı 1.100'ü aşan IDC analisti, dünyanın 110'dan fazla ülkesinde, teknoloji ve sektör fırsatları ve eğilimleri konusunda küresel, bölgesel ve yerel uzmanlık sağlar. IDC, 50 yıldır müşterilerin temel iş hedeflerine ulaşmasına yardımcı olmak amacıyla stratejik içgörüler sağlamaktadır. IDC, dünyanın lider teknoloji medya, araştırma ve etkinlik şirketi olan IDG'nin bir yan kuruluşudur.

Genel Merkezler

140 Kendrick Street
Building B
Needham, MA 02494
ABD
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com

Telif Hakkı Bildirimi

IDC Bilgilerinin ve Verilerinin Harici Olarak Yayımlanması – IDC'nin, reklam, basın bülteni ya da tanıtım malzemesinde kullanılacak her türlü bilgisi IDC'nin ilgili Başkan Yardımcısından ya da Ülke Müdüründen önceden yazılı onay alınmasını gerektirir. Bu tür bir onay isteği önerilen belge taslağıyla birlikte sunulmalıdır. IDC'nin herhangi bir nedenle harici kullanım onayı vermeme hakkı saklıdır.

Copyright 2021 IDC. Yazılı izin alınmadan çoğaltılması kesinlikle yasaktır.

