

### 3. 컴퓨팅 방식

#### 3.1. 인메모리 컴퓨팅 (In-Memory Computing): 데이터를 메모리에서 바로 처리하는 초고속 컴퓨팅

##### ① 왜 나오게 되었나?

기존 컴퓨팅 방식은 데이터를 저장하는 메모리와 데이터를 처리하는 프로세서가 분리되어 있어, 데이터를 주고받는 과정에서 병목 현상이 발생하고 처리 속도가 느려지는 문제가 있었다. 마치 서류를 보관하는 캐비닛과 서류를 처리하는 사무실이 멀리 떨어져 있어, 직원이 서류를 가져오고 다시 돌려놓는 데 많은 시간을 소모하는 것과 같은 상황이었다.

특히 빅데이터 시대에 데이터 처리량이 폭발적으로 증가하면서 이러한 문제는 더욱 심각해졌다. 마치 서류의 양이 급증하면서 직원이 서류를 옮기는 데만 대부분의 시간을 쓰게 되는 것처럼, 데이터를 메모리와 프로세서 사이에서 옮기는 데 많은 시간과 에너지가 낭비되었다. 이를 해결하기 위해 데이터를 메모리에서 바로 처리하는 인메모리 컴퓨팅 기술이 등장했다.

## ② 쉽게 설명하면?

인메모리 컴퓨팅은 데이터를 하드디스크나 SSD(Solid State Drive)와 같은 저장장치에 저장하지 않고, 메모리에 저장하여 처리하는 기술이다. 메모리는 저장장치보다 훨씬 빠른 속도로 데이터를 읽고 쓸 수 있기 때문에, 데이터 처리 속도를 획기적으로 높일 수 있다.

이는 마치 책상 위에 필요한 서류를 모두 펼쳐놓고 바로 일하는 것과 같다. 서류를 캐비닛에서 꺼내고 다시 돌려놓는 시간이 없기 때문에, 훨씬 빠르고 효율적으로 일을 처리할 수 있다. 인메모리 컴퓨팅은 이와 같이 데이터를 메모리에 저장하여 프로세서가 바로 접근하고 처리할 수 있게 함으로써, 데이터 처리 속도를 극대화한다.

## ③ 예시와 함께 좀 더 자세히 알아볼까?

인메모리 컴퓨팅은 실시간 데이터 분석, 금융 거래 처리, 인공지능 학습 등 대용량 데이터를 빠르게 처리해야 하는 분야에서 활용된다.

예를 들어, 주식 시장에서 실시간으로 주가 변동을 분석하고 투자 결정을 내리는 데 인메모리 컴퓨팅이 사용될 수 있다. 주가는 매초마다 변화하므로, 방대한 양의 실시간 데이터를 즉각적으로 처리해야 한다. 인메모리 컴퓨팅을 활용하면 주가 데이터를 메모리에 저장하고 바로 분석할 수 있어, 지연 시간 없이 신속한 투자 결정을 내릴 수 있다.

또한, 인공지능 모델 학습 시에도 인메모리 컴퓨팅이 활용된다. 인공지능 모델을 학습시키려면 대규모 데이터를 반복적으로 처리해야 하는데, 데이터를 메모리에 로드하여 고속으로 처리함으로써 학습 시간을 크게 단축할 수 있다. 이는 마치 방대한 자료를 책상 위에 펼쳐놓고 공부하는 것과 같다. 자료를 찾고 정리하는 시간을 줄이고 학습 효율을 높일 수 있는 것이다. 다만 이처럼 데이터를 메모리에 모두 펼쳐놓았을 때 학습 효율이 높아지는 것은 AI 반도체가 데이터를 하나씩 순차적으로 찾는 것이 아니라, 동시에 병렬적으로 처리할 수 있기 때문에 가능한 것이다.

## ④ 왜 중요한가?

인메모리 컴퓨팅은 AI 시대에 필수적인 기술이다. 데이터 처리 속도를 획기적으로 높여 실시간 데이터 분석, 빅데이터, 머신러닝 등 다양한 분야에서 혁신을 가져올 수 있다.

이는 마치 업무 속도를 높임으로써 더 많은 일을 처리하고, 더 나은 의사결정을 내릴 수 있는 것과 같다. 인메모리 컴퓨팅은 데이터 처리의 속도와 효율을 극대화함으로써, 기존에는 불가능했던 실시간 분석과 예측을 가능하게 한다.

또한, 데이터 처리 과정에서 발생하는 에너지 소비를 줄여 친환경적인 컴퓨팅 환경을 구축하는 데에도 기여한다. 데이터를 메모리와 프로세서 사이에서 옮기는 데 드는 에너지를 절약할 수 있기 때문이다. 이는 마치 서류를 옮기는 데 소모되는 시간과 노력을 줄임으로써, 더 효율적이고 지속가능한 업무 환경을 만드는 것과 같다.

#### ⑤ 어디에 활용되는가?

인메모리 컴퓨팅은 금융, 의료, 제조, 유통 등 다양한 산업 분야에서 활용된다. 예를 들어, 금융 분야에서는 실시간 사기 탐지, 고객 맞춤형 상품 추천 등에 활용된다. 수많은 금융 거래 데이터를 실시간으로 분석하여 이상 징후를 포착하거나, 고객의 행동 패턴을 분석하여 최적의 금융 상품을 제안할 수 있다.

의료 분야에서는 환자 데이터 분석, 질병 예측 등에 활용된다. 방대한 의료 기록과 검사 결과를 고속으로 분석하여 질병을 조기에 진단하고, 환자에게 최적의 치료 방법을 제안할 수 있다.

제조 분야에서는 생산 공정 최적화, 품질 관리 등에 활용된다. 공장에서 생산되는 제품의 데이터를 실시간으로 모니터링하여 불량품을 조기에 발견하고, 생산 효율을 극대화할 수 있다.

유통 분야에서는 재고 관리, 판매 예측 등에 활용된다. 판매 데이터를 분석하여 상품 수요를 예측하고, 적절한 재고를 유지함으로써 판매 기회 손실을 최소화하고 수익을 극대화할 수 있다.

이처럼 인메모리 컴퓨팅은 데이터 처리 속도와 효율이 중요한 다양한 분야에서 핵심적인 역할을 하고 있다. 앞으로도 인메모리 컴퓨팅 기술의 발전은 산업 전반의 디지털 혁신을 가속화할 것으로 기대된다.

### 3.2. 희소 컴퓨팅 (Sparse Computing): 0을 효율적으로 활용하는 똑똑한 계산법

#### ① 왜 나오게 되었나?

AI 모델은 점점 더 복잡해지고 거대해지면서, 모델의 학습과 실행에 필요한 연산량과 메모리 사용량이 폭발적으로 증가했다. 그러나 대부분의 데이터는 0을 포함한 의미 없는 정보를 많이 포함하고 있다. 마치 책에서 내용과 관련 없는 빈 페이지가 대부분을 차지하는 것과 같은 상황이었다.

이러한 비효율성은 AI 모델의 학습과 실행 속도를 저하시키고, 막대한 컴퓨팅 자원을 낭비하는 결과를 초래했다. 한정된 자원을 효율적으로 활용하고, 더욱 복잡한 인공지능 모델을 다루기 위해서는 새로운 접근 방식이 필요했다. 이에 따라 데이터의 희소성(sparsity)을 활용하여 연산 효율을 극대화하는 희소 컴퓨팅 기술이 등장하게 되었다.

## ② 쉽게 설명하면?

희소 컴퓨팅은 데이터에서 0이나 의미 없는 값을 제외하고, 유의미한 값만을 이용하여 연산을 수행하는 기술이다. 마치 빈칸이 많은 시험지를 채점할 때 빈칸은 무시하고 정답만 확인하는 것과 비슷하다.

이해를 돕기 위해 책의 예시를 다시 들어보자. 500페이지 중 100페이지만 실제 내용이 있는 책이 있다고 가정해보자. 기존의 방식대로라면 500페이지 전체를 다 읽어야 하지만, 희소 컴퓨팅을 적용하면 내용이 있는 100페이지만 선별적으로 읽으면 된다. 이를 통해 책을 읽는 데 걸리는 시간과 노력을 크게 줄일 수 있다. 데이터 처리에서도 이와 같은 원리가 적용된다. 0이나 의미 없는 값을 건너뛰고 유의미한 값만 처리하므로, 연산량과 메모리 사용량을 줄이고 AI 모델의 학습 및 실행 속도를 높일 수 있다.

## ③ 예시와 함께 좀 더 자세히 알아볼까?

자연어 처리에서 사용되는 단어 임베딩(word embedding)을 예로 들어보자. 단어 임베딩은 각 단어를 수백 개의 숫자로 이루어진 벡터로 표현하는데, 대부분의 벡터는 0을 많이 포함하고 있다.

예를 들어, “고양이”라는 단어의 임베딩 벡터가 [0, 0, 0.5, 0, 0.2, 0, ..., 0.8] 이라고 해보자. 대부분의 원소가 0이고, 의미 있는 값은 몇 개뿐이다. 희소 컴퓨팅은 이러한 0을 모두 저장하지 않고, [3: 0.5, 5: 0.2, ..., 300: 0.8]와 같이 의미 있는 값의 위치와 값만 저장한다. 이렇게 하면 메모리 사용량을 크게 줄일 수 있다. 또한 연산 시에도 0이 아닌 값만 고려하면 되므로, 연산량도 크게 감소한다.

이는 마치 책에서 중요한 내용만 발췌하여 요약본을 만드는 것과 같다. 중요하지 않은 내용은 과감히 생략하고, 핵심 내용만 간추려 담으면 책의 분량을 줄일 수 있고, 읽는 시간도 단축할 수 있다. 희소 컴퓨팅은 데이터의 요약본을 만들어 처리하는 것이라고 이해할 수 있다.

## ④ 왜 중요한가?

희소 컴퓨팅은 인공지능 모델의 효율성을 획기적으로 높이는 기술이다. 연산량과 메모리 사용량을 줄여 AI 모델의 학습 및 실행 속도를 높이고, 더욱 복잡하고 거대한 모델을 효율적으로 처리할 수 있게 한다.

이는 마치 책을 더 빨리 읽고 이해할 수 있게 해주는 속독법과 같다. 희소 컴퓨팅을 통해 AI 모델은 방대한 데이터를 더욱 빠르게 학습하고, 실시간으로 추론을 수행할 수 있게 된다. 이는 인공지능 기술의 실용성을 높이고, 더 다양한 분야에서 인공지능을 활용할 수 있는 토대가 된다.

또한 희소 컴퓨팅은 인공지능 모델의 에너지 효율성을 높이는 데에도 기여한다. 불필요한 연산을 줄임으로써 전력 소모를 감소시키고, 더 친환경적인 AI 시스템을 구축할 수 있게 된다. 이는 AI의 지속가능성을 높이고, 더 넓은 범위에서 AI기술을 활용할 수 있는 발판이 된다.