|  |  |
| --- | --- |
| 1. 2) Multicloud Deployment of AI Workflows Using FaaS and Storage Services | |
| **저자** | D. Chahal, S. C. Palepu, R. Singhal |
| **학회** | HiPS '22 (2nd Workshop on High Performance Serverless Computing) |
| **한줄 요약** | |
| 내용  : FaaS(Functio-as-a-Service)와 클라우드 스토리지 서비스를 사용하여 AI 추론 워크플로우를 멀티클라우드 환경에 배포하는 과정에 대한 실증적 연구를 제시함. | |
| **요약** | |
| 내용  : 다중 클라우드 환경에서 AI 워크플로우를 효율적으로 배포하기 위해 FaaS 플랫폼과 스토리지 서비스를 결합하는 방법을 연구하고, 비용과 성능을 분석하며, 다중 클라우드 배포의 장점을 제시하고 비용 모델을 제안하는 내용을 다룸. | |
| **강점/약점** | |
| 내용   * 강점 * 멀티클라우드 환경의 최적화: 이 논문은 멀티클라우드 환경에서 AI 워크플로우를 효율적으로 배포하는 방법을 제시함으로써, 다양한 클라우드 서비스 간의 최적화된 리소스 활용 방안을 탐색함. * 비용 효율성: FaaS와 클라우드 스토리지 서비스를 활용하여 AI 워크플로우를 배포함으로써, 비용 효율성을 높이는 방법을 제시함. 특히, 데이터 전송 비용 절감에 초점을 맞추어 클라우드 기반 AI 애플리케이션의 전반적인 비용을 최적화할 수 있음. * 약점 * 멀티클라우드 복잡성: 멀티클라우드 환경은 다양한 클라우드 서비스 제공업체의 API와 서비스 특성을 이해하고 통합해야 하는 복잡성을 내포하고 있음. 이 논문에서 제시한 방법이 이러한 복잡성을 완전히 해결하지 못할 수 있음. * 일반화의 한계: 특정 FaaS와 스토리지 서비스에 초점을 맞춘 연구 결과는 다른 클라우드 서비스 제공업체의 서비스에 적용할 때 일반화하는 데 한계가 있을 수 있음. | |
| **Minor Comments** | |
| 오타 / 논리 부족 / 용어 미흡  : 논문을 검토하는 과정에서 위와 같은 부분은 없었음. | |