

1. 우아함 (우승자: 파이썬)

우아함에 대한 평가는 매우 주관적이다. 맷로프 교수는 파이썬은 괄호나 교정기 사용을 크게 줄일 수 있기 때문에 파이썬의 명백한 승리이며, 파이썬은 매우 낯설하다고 말했다.

2. 학습 곡선 (우승자: R)

Python을 사용하기 위해선 NumPy, Pandas, matplotlib를 포함한 많은 자료를 배워야 한다. 하지만 R에는 기본적으로 매트릭스 유형과 기본 그래픽이 내장되어 있다.

시스템에 능숙한 경우에도 Python의 라이브러리를 구성하기 까다로울 수 있다. 반면 대부분의 R 패키지는 즉시 실행되며, 초보자도 R을 활용해 몇 분 안에 간단한 데이터 분석을 할 수 있다.

3. 사용 가능한 라이브러리 (우승부)

파이썬 패키지 인덱스(PyPI)는 183,000개 이상의 패키지를 가지고 있으며, CRAN(Compreged Rarchive Network)은 12,000개 이상의 패키지를 가지고 있다. 그러나 데이터 과학에 대한 PyPI는 적다.

예를 들어 맷로프가 주어진 데이터 포인트의 가장 가까운 이웃에 대한 빠른 계산을 하기 위해 코드가 필요할 때, 두 개의 패키지를 즉시 찾을 수 있었다. 반면 파이썬에서는 가장 가까운 이웃의 코드를 찾을 수 없었다. 그는 간단하고 직설적인 구현은 오직 한 개였다고 언급했다.

또한 맷로프는 PyPI에 다음 용어를 검색하면 아무것도 나타나지 않았다고 덧붙였다: 로그 선형 모델, 포아송 회귀 분석, 기약 변수, 공간 데이터, FWE 비율.

4. 머신 러닝 (우승자: Python, 하지만 많이는 아님)

최근 몇 년간 파이썬의 엄청난 성장은 머신러닝과 인공지능의 상승에 일부분 영향을 받았다. Python은 AlexNet과 같이 이미지 인식을 위해 정밀하게 조정된 라이브러리를 많이 제공하지만, R 버전도 쉽게 개발할 수 있다.

맷로프는 파이썬 라이브러리의 힘은 특정한 이미지 평탄화 ops 설정에 의해 생기며, R의 Keras wrapper를 통해 쉽게 구현할 수 있다고 말했다. 순수한 R 버전의 TensorFlow(텐서 플로우)가 개발될 수 있으며, R의 랜덤 포레스트와 그래디언트 부스팅의 가용성이 탁월할 수 있다고 주장한다.

5. 통계적 정확성 (우승자: R, 월등히 좋음)

맷로프는 파이썬을 옹호하는 머신러닝 전문가들이 때때로 관련된 통계적 문제에 대해 잘 이해하지 못한다고 지적했다. 반면 R은 통계학자에 의해 작성되었다고 덧붙였다.

6. 병렬 연산 (우승부)

R과 Python의 기본 버전은 멀티 코어 연산에 대한 지원이 약하다. 파이썬의 멀티프로세싱 패키지는 다른 이슈들에 대한 좋은 해결책이 아니며, R의 병렬 패키지도 마찬가지이다.

맷로프는 "클러스터 연산을 지원하는 외부 라이브러리는 두 언어 모두 괜찮다", "현재 Python은 GPU에 대한 더 나은 인터페이스를 가지고 있다."라고 말했다.

7. C/C++ 인터페이스 (우승자: R, 하지만 많이는 아님)

R의 Rcpp는 R과 C/C++를 연결하는 강력한 도구이다. 파이썬 swig(스위그)와 같은 도구를 가지고 있지만 그것은 그렇게 강력하지 않으며, Pybind11 패키지는 아직 개발 중이다. R의 새로운 ALTREP 아이디어는 성능과 사용성을 향상시킬 수 있는 잠재력을 가지고 있다. 하지만 Cython과 파이썬의 PyPy 변형은 때때로 C/C++ 인터페이스의 필요성을 없앨 수 있다.

8. 객체 방향, 메타 프로그래밍 (우승자: R, 하지만 많이는 아님)

R과 Python 둘 다에서 기능이 객체임에도 불구하고 R은 이를 더욱 심각하게 받아들이고 있다고 언급했다. 그는 "파이썬을 사용할 때마다 R에서 많이 하는 것처럼 단말기에 기능을 인쇄하지 못한다는 사실에 짜증이 난다"고 말했다.

Python은 단지 하나의 OOP 패러다임을 가지고 있다. 반면 R에서는 몇 가지를 선택할 수 있다. 그는 R의 마법 메타 프로그래밍 기능(코드를 생성하는 코드)을 감안할 때 컴퓨터 과학자들은 R에 대해 주목해야 한다고 주장한다.

9. 언어의 통일성 (우승자: 파이썬, 월등히 좋음)

파이썬이 버전 2.7에서 3.x로 변해도 큰 혼란이 생기지 않을 것이다. 하지만 R은 RStudio의 영향으로 두 개의 다른 언어로 바뀌어가고 있다. (R, Tidyverse)

맷로프는 Tidyverse가 일반적인 R보다 우월하다면 받아들일 수 있지만, 개인적으로 그렇게 생각하지 않는다고 언급했다. 또한 이것이 초보자들에게 어렵게 만든다고 말했다.

10. 링크된 데이터 구조 (우승자: 파이썬)

Binary tree 같은 고전적인 컴퓨터 사이언스 데이터 구조는 파이썬에서 구현하기 쉽다. R에서 리스트 클래스를 사용해 구현할 수 있지만, 맷로프는 이것고 느리다고 추측한다.