

<알고리즘 선택 시 고려 사항>

알고리즘을 선택할 때에는 언제나 정확성, 학습 시간, 사용 편의성을 고려해야 한다.

많은 경우 정확성을 최우선으로 두는데. 반면 초급자는 가장 잘 알고 있는 알고리즘에 초점을 맞추는 경향이 있다.

데이터 세트가 제공됐을 때 가장 먼저 고려해야 할 것은 ‘어떤 결과가 나올 것인지에 상관없이 어떻게 결과를 얻을 것인가’이다. 초급자일수록 실행하기 쉽고 결과를 빨리 얻을 수 있는 알고리즘을 선택하기 쉬운데. 프로세스의 첫 단계에서는 관찰을 수 있겠지만 일부 결과를 얻었고 데이터에 익숙해진 후라면 정교한 알고리즘을 사용하는 데 시간을 더 많이 할애해야 한다. 그래야만 데이터를 더욱 잘 이해하고, 결과를 개선시킬 수 있다.

심지어 이 단계에서조차도 최상의 알고리즘은 가장 높은 정확성을 달성한 방법이 아닐 수 있다. 일반적으로 알고리즘은 달성 가능한 최고의 성능을 발휘하기 위해 세심한 튜닝(tuning)과 광범위한 학습을 요구하기 때문이다.