



유추 파이프라인 만들기

8분

파이프라인을 만들고 실행해 클러스터링 모델을 학습한 후에는 모델을 사용하여 클러스터에 새 데이터 관찰을 할당하는 추론 파이프라인을 만들 수 있습니다. 이를 통해 애플리케이션에서 사용하도록 게시할 수 있는 예측 서비스의 기반이 형성됩니다.

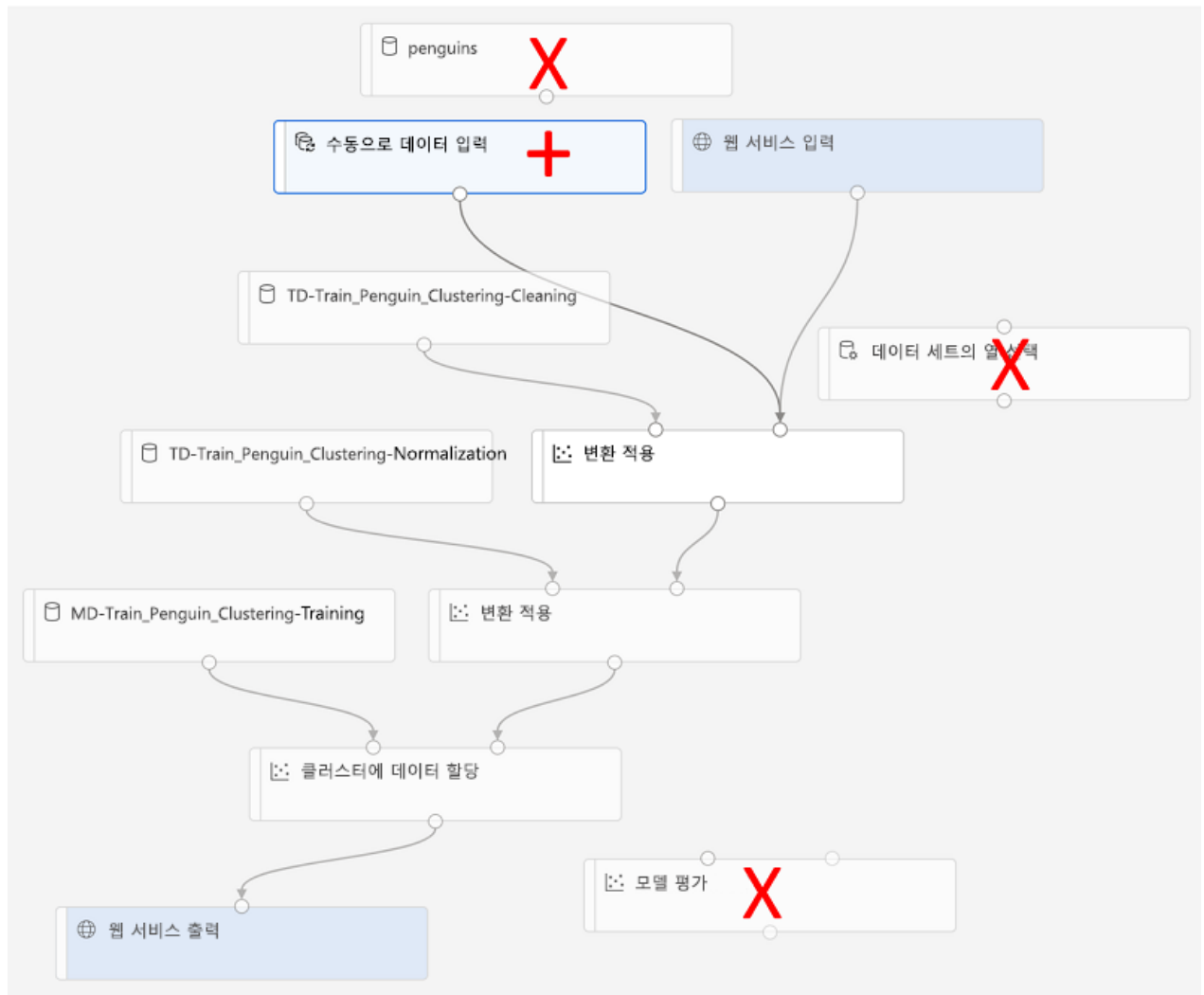
유추 파이프라인 만들기

1. Azure Machine Learning Studio에서 이전에 만든 **Train Penguin Clustering** 파이프라인을 엽니다.
2. **유추 파이프라인 만들기** 드롭다운 목록에서 **실시간 유추 파이프라인** 을 클릭합니다. 몇 초 후에 **Train Penguin Clustering-real time inference** 라는 새 버전의 파이프라인이 열립니다.

파이프라인에 **웹 서비스 입력** 과 **웹 서비스 출력** 모듈이 포함되지 않은 경우 **디자이너** 페이지로 돌아가 **Train Penguin Clustering-real time inference** 파이프라인을 다시 엽니다.

3. 새 파이프라인의 이름을 **Predict Penguin Clusters** 로 바꾸고 새 파이프라인을 검토합니다. 여기에는 제출할 새 데이터에 대한 웹 서비스 입력과 결과를 반환하기 위한 웹 서비스 출력이 포함되어 있습니다. 학습 파이프라인의 변환 및 클러스터링 모델은 학습 데이터의 통계를 기반으로 이 파이프라인에 캡슐화되며, 새 데이터의 변환 및 채점에 사용됩니다.

유추 파이프라인을 다음과 같이 변경합니다.



- **penguin-data** 데이터 세트를 **Species** 열을 포함하지 않은 **수동으로 데이터 입력** 모듈로 바꿉니다.
- 이제는 중복되는 **데이터 세트에서 열 선택** 모듈을 제거합니다.
- **웹 서비스 입력** 및 **수동으로 데이터 입력** 모듈(클러스터링될 데이터의 입력을 나타냄)을 **변환 적용** 모듈에 연결합니다.
- **모델 평가** 모듈을 제거합니다.

파이프라인을 수정할 때 위 이미지와 정보를 참조로 사용하여 아래의 나머지 단계를 수행합니다.

- 추론 파이프라인은 새 데이터가 원래 학습 데이터의 스키마와 일치하는 것으로 가정하므로 학습 파이프라인의 **penguin-data** 데이터 세트가 포함됩니다. 하지만 이 입력 데이터에는 모델에서 사용하지 않는 펭귄 종에 대한 열이 포함되어 있습니다. **penguin-data** 데이터 세트와 **데이터 세트에서 열 선택** 모듈을 모두 삭제하고 **데이터 입력 및 출력** 섹션에 있는 **수동으로 데이터 입력** 모듈로 대체합니다. 그런 다음 **수동으로 데이터 입력** 모듈의 설정을 수정하여 다음 CSV 입력을 사용합니다. 여기에는 3개의 새로운 펭귄 관찰에 대한 특징 값이 포함되어 있습니다(헤더 포함).

CSV

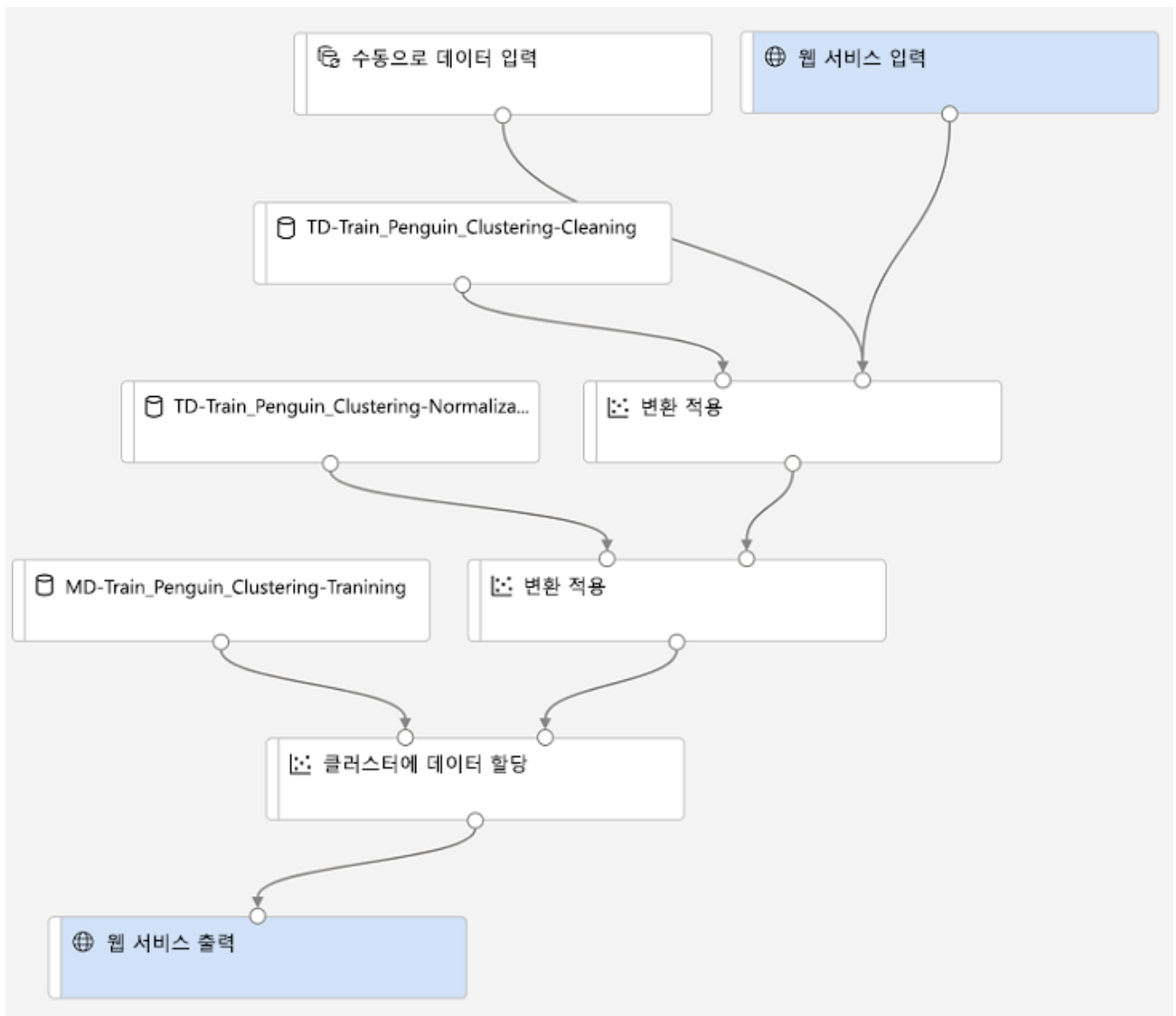
복사

```

CulmenLength,CulmenDepth,FlipperLength,BodyMass
39.1,18.7,181,3750
49.1,14.8,220,5150
46.6,17.8,193,3800

```

5. 웹 서비스 입력 및 수동으로 데이터 입력 모듈의 출력을 첫 번째 변환 적용 모듈의 데이터 세트(오른쪽) 입력에 연결합니다.
6. 모델 평가 모듈을 삭제합니다.
7. 파이프라인이 다음과 유사한지 확인합니다.



8. 컴퓨팅 클러스터에서 **mslearn-penguin-inference** 라는 새 실험으로 파이프라인을 제출합니다. 이 작업은 다소 시간이 걸릴 수 있습니다.
9. 파이프라인이 완료되면 **클러스터에 데이터 할당** 모듈의 **결과 데이터 세트** 출력을 시각화하여 입력 데이터의 3가지 펭귄 관찰에 대한 예측된 클러스터 할당 및 메트릭을 확인합니다.

추론 파이프라인은 특징을 기반으로 펭귄 관찰을 클러스터에 할당합니다. 이제 클라이언트 애플리케이션에서 사용할 수 있도록 파이프라인을 게시할 준비가 되었습니다.

다음 단원: 예측 서비스 배포

[계속 >](#)