**Membuat Model Pengklusteran dengan perancang Azure Machine Learning**

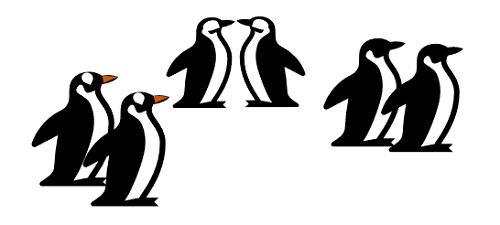
Pengklusteran adalah teknik pembelajaran mesin tanpa pengawasan yang digunakan untuk mengelompokkan entitas serupa berdasarkan fiturnya. Pelajari cara membuat model pengklusteran menggunakan perancang Azure Machine Learning.

**Tujuan pembelajaran**

Melatih dan menerbitkan model pengklusteran dengan perancang Azure Machine Learning

1. **Pengantar**

Pengklusteran adalah bentuk pembelajaran mesin yang digunakan untuk mengelompokkan item serupa ke dalam kluster berdasarkan fiturnya. Misalnya, seorang peneliti mungkin mengambil pengukuran penguin, dan mengelompokkannya berdasarkan kesamaan dalam proporsinya.



Pengklusteran adalah contoh pembelajaran mesin tanpa pengawasan, tempat Anda melatih model untuk memisahkan item ke dalam kluster berdasarkan murni pada karakteristiknya, atau fitur. Tidak ada nilai kluster (atau label) yang diketahui sebelumnya untuk melatih model.

Anda dapat menggunakan perancang Microsoft Azure Machine Learning untuk membuat model pengklusteran dengan menggunakan antarmuka visual drag dan drop, tanpa perlu menulis kode apa pun.

Dalam modul ini, Anda akan mempelajari cara:

* Gunakan perancang Azure Machine Learning untuk melatih model pengklusteran.
* Gunakan model pengklusteran untuk inferensi.
* Sebarkan model pengklusteran sebagai layanan.

Untuk menyelesaikan modul ini, Anda memerlukan langganan Microsoft Azure. Jika belum memilikinya, Anda dapat mendaftar untuk uji coba gratis di <https://azure.microsoft.com.>

1. **Membuat ruang kerja Azure Machine Learning**

Azure Machine Learning adalah platform berbasis cloud untuk membuat dan mengoperasikan solusi pembelajaran mesin di Azure. Ini termasuk berbagai fitur dan kemampuan yang membantu ilmuwan data menyiapkan data, melatih model, menerbitkan layanan prediktif, dan memantau penggunaannya. Salah satu fitur ini adalah antarmuka visual yang disebut perancang, yang dapat Anda gunakan untuk melatih, menguji, dan menyebarkan model pembelajaran mesin tanpa menulis kode apa pun.

1. **Membuat ruang kerja Azure Machine Learning**

Untuk menggunakan Azure Machine Learning, Anda membuat ruang kerja di langganan Azure. Kemudian, Anda dapat menggunakan ruang kerja ini untuk mengelola data, sumber daya komputasi, kode, model, dan artefak lainnya yang terkait dengan beban kerja pembelajaran mesin Anda.

**Catatan** :

Modul ini adalah salah satu dari banyak modul yang menggunakan ruang kerja Azure Machine Learning, termasuk modul lain di jalur pembelajaran **Dasar-Dasar Microsoft Azure AI: Menjelajahi alat visual untuk pembelajaran Mesin : https://learn.microsoft.com/id-id/training/paths/create-no-code-predictive-models-azure-machine-learning/** . Jika menggunakan langganan Azure Anda sendiri, Anda dapat mempertimbangkan untuk membuat ruang kerja sekali dan menggunakannya kembali di modul lain. Langganan Azure Anda akan dikenakan biaya kecil untuk penyimpanan data selama ruang kerja Azure Machine Learning ada di langganan Anda, jadi sebaiknya hapus ruang kerja Azure Machine Learning saat tidak lagi diperlukan.

Jika Anda belum memilikinya, ikuti langkah-langkah ini untuk membuat ruang kerja:

1. Masuk ke portal Azure menggunakan info masuk Microsoft Anda.
2. Pilih Buat sumber daya, cari Pembelajaran Mesin, dan buat sumber daya Azure Machine Learning baru dengan paket Azure Machine Learning. Gunakan pengaturan berikut:

* **Langganan**: Langganan Azure Anda
* **Grup Sumber Daya**: Buat atau pilih grup sumber daya
* **Nama ruang kerja**: Masukkan nama unik untuk ruang kerja Anda
* **Wilayah**: Pilih wilayah geografis terdekat
* **Akun penyimpanan**: Perhatikan akun penyimpanan baru default yang akan dibuat untuk ruang kerja Anda
* **Brankas kunci:** Perhatikan brankas kunci baru default yang akan dibuat untuk ruang kerja Anda
* **Wawasan aplikasi:** Perhatikan sumber daya wawasan aplikasi baru default yang akan dibuat untuk ruang kerja Anda
* **Registri kontainer:** Tidak ada (satu registri kontainer akan dibuat secara otomatis saat pertama kali Anda menyebarkan model ke kontainer)

1. Pilih **Tinjau + buat**. Tunggu hingga ruang kerja Anda dibuat (bisa memakan waktu beberapa menit). Kemudian, buka di portal.
2. Pada halaman **Gambaran Umum** untuk ruang kerja Anda, luncurkan studio Azure Machine Learning (atau buka tab browser baru dan navigasikan ke https://ml.azure.com), dan masuk ke studio Azure Machine Learning menggunakan akun Microsoft Anda.
3. Dalam studio Azure Machine Learning, pilih tiga ikon baris di bagian kiri atas untuk menampilkan berbagai halaman di antarmuka. Anda dapat menggunakan halaman ini untuk mengelola sumber daya di ruang kerja.

Anda dapat mengelola ruang kerja menggunakan portal Azure, tetapi untuk ilmuwan data dan teknisi operasi Machine Learning, studio Azure Machine Learning menyediakan antarmuka pengguna yang lebih terfokus untuk mengelola sumber daya ruang kerja.

1. **Membuat sumber daya komputasi**

Untuk melatih dan menyebarkan model menggunakan perancang Azure Machine Learning, Anda memerlukan target komputasi untuk menjalankan proses pelatihan. Anda juga akan menggunakan target komputasi ini untuk menguji model terlatih setelah penyebaran.

1. **Membuat target komputasi**

Target komputasi adalah sumber daya berbasis cloud tempat Anda dapat menjalankan pelatihan model dan proses eksplorasi data.

Di Azure Machine Learning studio : https://ml.azure.com/, perluas panel kiri dengan memilih tiga baris di kiri atas layar. Lihat halaman Komputasi (di bawah Kelola). Anda mengelola target komputasi untuk aktivitas ilmu data Anda di studio. Ada empat jenis sumber daya komputasi yang dapat Anda buat:

* **Instans Komputasi**: Stasiun kerja pengembangan yang dapat digunakan oleh ilmuwan data untuk bekerja dengan data dan model.
* **Kluster Komputasi:** Kluster komputer virtual yang dapat diskalakan untuk pemrosesan kode eksperimen sesuai permintaan.
* **Kluster Inferensi:** Target penyebaran untuk layanan prediktif yang menggunakan model terlatih Anda.
* **Komputasi Terlampir:** Tautan ke sumber daya komputasi Azure yang ada, seperti kluster Virtual Machines atau Azure Databricks.

**Catatan**:

Instans dan kluster komputasi didasarkan pada gambar komputer virtual Azure standar. Untuk modul ini, gambar Standard\_DS11\_v2 disarankan untuk mencapai keseimbangan biaya dan performa yang optimal. Jika langganan Anda memiliki kuota yang tidak menyertakan gambar ini, pilih gambar alternatif; tetapi perlu diingat bahwa gambar yang lebih besar dapat dikenakan biaya yang lebih tinggi dan gambar yang lebih kecil mungkin tidak cukup untuk menyelesaikan tugas. Atau, minta administrator Azure Anda memperpanjang kuota Anda.

1. Buka tab Kluster Komputasi, dan tambahkan kluster komputasi baru dengan pengaturan berikut:

* **Lokasi**: Pilih yang sama dengan ruang kerja Anda. Jika lokasi tidak tercantum, pilih yang terdekat dengan Anda
* **Tingkat Mesin Virtual:** Khusus
* **Jenis Komputer Virtual**: CPU
* **Ukuran Komputer Virtual:**
* Pilih opsi Pilih dari semua opsi
* Cari dan pilih Standard\_DS11\_v2

1. Pilih **Selanjutnya**

* **Nama komputasi**: masukkan nama unik
* **Jumlah minimum node**: 0
* **Jumlah maksimum node**: 2
* **Detik menganggur sebelum menurunkan skala**: 120
* **Aktifkan akses SSH**: Tidak dipilih

1. Pilih **Buat**

**Tip :**

Setelah Anda menyelesaikan seluruh modul, pastikan untuk mengikuti petunjuk Bersihkan di akhir modul untuk menghentikan sumber daya komputasi Anda. Hentikan sumber daya komputasi Anda untuk memastikan langganan Anda tidak dikenakan biaya.

**4. Jelajahi data**

Untuk melatih model pengklusteran, Anda memerlukan himpunan data yang mencakup beberapa pengamatan item yang ingin Anda klusterkan, termasuk fitur numerik yang dapat digunakan untuk menentukan kesamaan antara kasus individual yang akan membantu memisahkannya menjadi kluster.

1. **Buat himpunan data**

Di Azure Machine Learning, data untuk pelatihan model dan operasi lainnya biasanya dienkapsulasi dalam objek yang disebut himpunan data. Dalam modul ini, Anda akan menggunakan himpunan data yang mencakup pengamatan tiga spesies penguin.

1. Di Azure Machine Learning studio, perluas panel kiri dengan memilih tiga baris di kiri atas layar. Lihat halaman Data (di bagian Aset). Halaman Data berisi file atau tabel data tertentu yang Anda rencanakan untuk digunakan di Azure ML. Anda juga dapat membuat himpunan data dari halaman ini.
2. Buat himpunan data dari file web, menggunakan pengaturan berikut:

* **Info dasar:**
* **URL Web**: https://aka.ms/penguin-data
* **Nama**: penguin-data
* **Jenis himpunan data:** Tabular
* **Deskripsi:** Data penguin
* **Lewati validasi data:** jangan pilih
* **Pengaturan dan pratinjau**:
* **Format file:** Dibatasi
* **Pemisah:** Koma
* **Pengodean**: UTF-8
* **Header kolom:** Hanya file pertama yang memiliki header
* **Lewati baris**: Tidak ada
* **Himpunan data berisi data multi-baris**: jangan pilih
* **Skema:**
* Sertakan semua kolom selain **Jalur**
* Meninjau jenis yang terdeteksi secara otomatis
* **Konfirmasikan detail:**
* Jangan membuat profil himpunan data setelah pembuatan

1. Setelah himpunan data dibuat, buka dan tampilkan halaman **Jelajahi** untuk melihat sampel data. Data ini menunjukkan pengukuran panjang dan kedalaman culmen (tagihan), panjang sirip, dan massa tubuh untuk beberapa pengamatan penguin. Ada tiga spesies penguin yang diwakili dalam himpunan data: Adelie, Gentoo, dan Chinstrap.

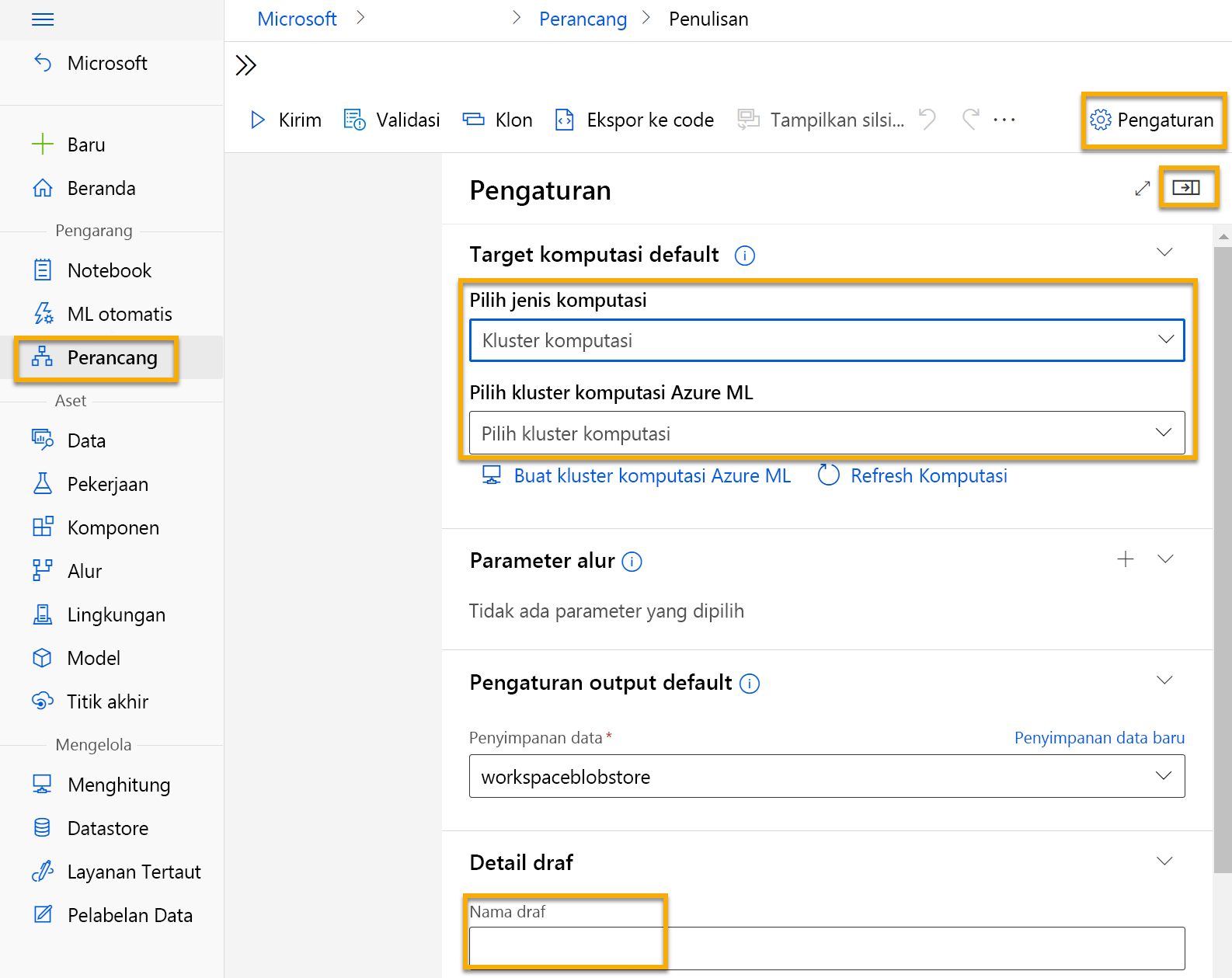
**Catatan :**

Himpunan data penguin yang digunakan dalam latihan ini adalah subset data yang dikumpulkan dan disediakan oleh Dr. Kristen Gorman dan Stasiun Palmer, Antartika LTER, anggota Jaringan Penelitian Ekologi Jangka Panjang.

1. **Membuat alur**

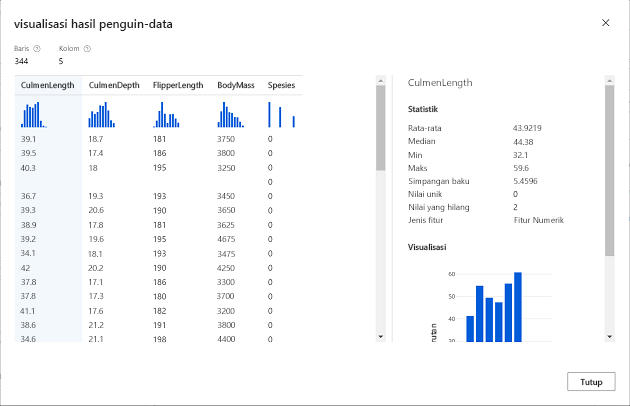
Untuk memulai dengan perancang pembelajaran mesin Azure, pertama-tama Anda harus membuat alur dan menambahkan himpunan data yang ingin Anda gunakan.

1. Di Azure Machine Learning studio, perluas panel kiri dengan memilih ikon tiga baris di bagian kiri atas layar. Lihat halaman **Perancang** (di bagian **Pembuat**), dan pilih tanda plus untuk membuat alur baru.
2. Di sisi kanan atas layar, pilih **Pengaturan**. Jika panel **Pengaturan** tidak terlihat, pilih ikon roda di samping nama saluran di bagian atas.
3. Di **Pengaturan**, Anda harus menentukan target komputasi untuk menjalankan alur. Di bagian Pilih **jenis komputasi,** pilih **Kluster komputasi.** Kemudian, pada **Pilih jenis komputasi Azure Machine Learning**, pilih kluster komputasi yang Anda buat sebelumnya.
4. Di **Pengaturan**, di bagian **Detail Draf**, ubah nama draf **(Pipeline-Created-on-date)** menjadi Latih **Pengklusteran Penguin**
5. Pilih ikon tutup di kanan atas panel **Pengaturan** untuk menutup panel.



1. **Memuat data ke kanvas**
2. Di sebelah nama saluran di sebelah kiri, pilih ikon panah untuk memperluas panel jika belum diperluas. Panel akan terbuka secara default ke panel **Pustaka Aset**, yang ditunjukkan oleh ikon buku di bagian atas panel. Perhatikan bahwa terdapat bilah pencarian untuk menemukan aset. Perhatikan dua tombol, **Data** dan **Komponen**.

1. Klik **Data**. Cari dan tempatkan himpunan data **penguin-data** ke kanvas.
2. Klik kanan (Ctrl+klik pada Mac) himpunan data **penguin-data** di kanvas, dan klik **Pratinjau data**.
3. Tinjau skema Profil data, yang menyatakan bahwa Anda dapat melihat distribusi berbagai kolom sebagai histogram. Lalu pilih kolom **CulmenLength**. Himpunan datanya akan terlihat mirip dengan ini:



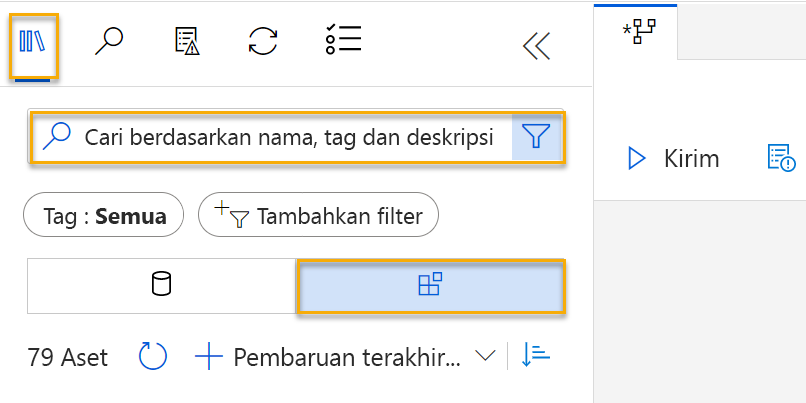
1. Perhatikan karakteristik himpunan data berikut:

* Himpunan data menyertakan kolom berikut:
* **CulmenLength**: Panjang tagihan penguin dalam milimeter.
* **CulmenDepth**: Kedalaman tagihan penguin dalam milimeter.
* **FlipperLength**: Panjang sirip penguin dalam milimeter.
* **BodyMass**: Berat penguin dalam gram.
* **Spesies**: Indikator spesies (0:"Adelie", 1:"Gentoo", 2:"Chinstrap")
* Ada dua nilai yang hilang di kolom **CulmenLength** (kolom **CulmenDepth**, **FlipperLength**, dan **BodyMass** juga memiliki dua nilai yang hilang).
* Nilai pengukuran dalam skala yang berbeda (dari puluhan milimeter hingga ribuan gram).

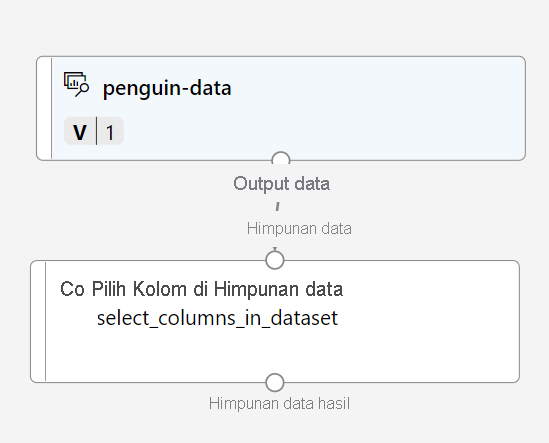
1. Tutup visualisasi himpunan data sehingga Anda dapat melihat himpunan data di kanvas alur pipa.
2. **Menerapkan transformasi**

Untuk mengelompokkan pengamatan penguin, kita hanya akan menggunakan pengukuran; jadi kita akan membuang kolom spesies. Kita juga perlu menghapus baris tempat nilai hilang, dan menormalkan nilai pengukuran numerik sehingga berada pada skala yang sama.

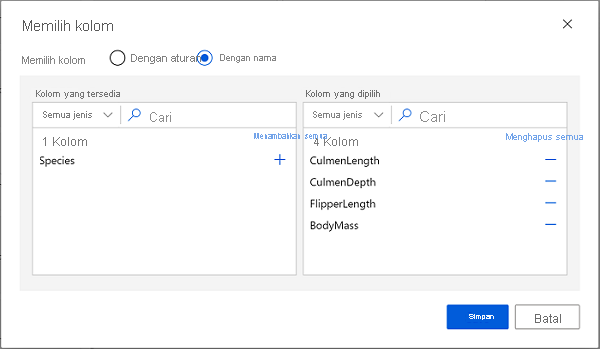
1. Di panel **Pustaka Aset** di sebelah kiri, klik **Komponen**, yang berisi berbagai modul yang dapat Anda gunakan untuk transformasi data dan pelatihan model. Anda juga dapat menggunakan bilah pencarian untuk menemukan modul dengan cepat.



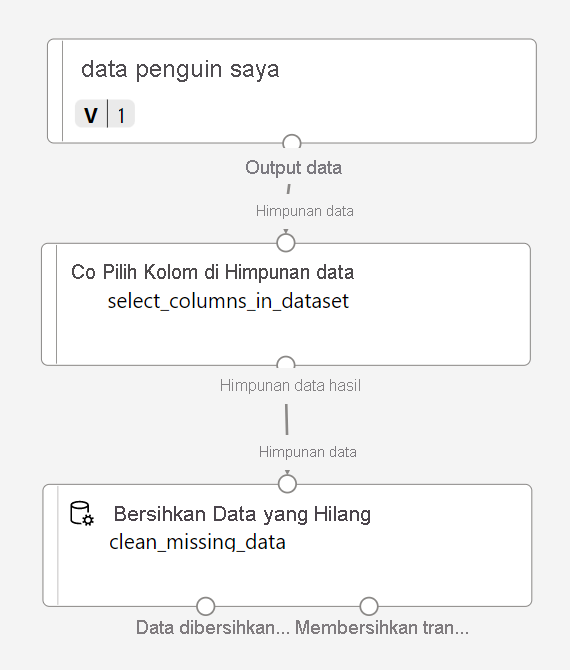
1. Untuk mengklusterkan pengamatan penguin, kita hanya akan menggunakan pengukuran - kita akan mengabaikan kolom spesies. Jadi, cari modul **Memilih Kolom dalam Himpunan Data** dan tempatkan ke kanvas, di bawah modul **penguin-data** dan sambungkan output di bagian bawah **penguin-data** ke input di bagian atas modul **Memilih Kolom dalam Himpunan Data**, seperti ini:



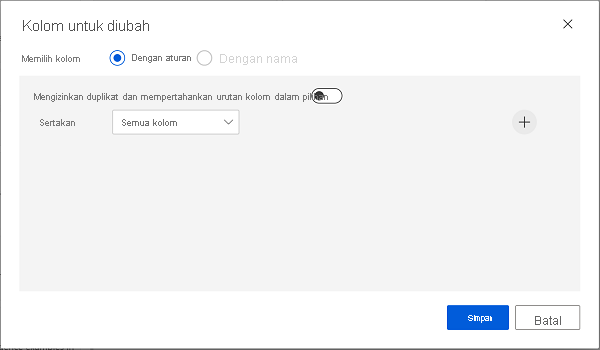
1. Pilih modul **Memilih Kolom dalam Himpunan Data**, dan pilih **Edit kolom** di panel sebelah kanan. Kemudian di jendela **Pilih kolom**, pilih **Berdasarkan nama** dan gunakan tautan untuk memilih nama kolom **CulmenLength**, **CulmenDepth**, **FlipperLength**, dan **BodyMass**; seperti ini:



1. Tutup pengaturan modul **Pilih Kolom dalam Himpunan Data** untuk kembali ke kanvas perancang.
2. Di **Pustaka aset**, cari modul **Membersihkan Data yang Hilang** dan tempatkan ke kanvas, di bawah modul **Memilih kolom dalam himpunan data** dan menyambungkannya seperti ini:



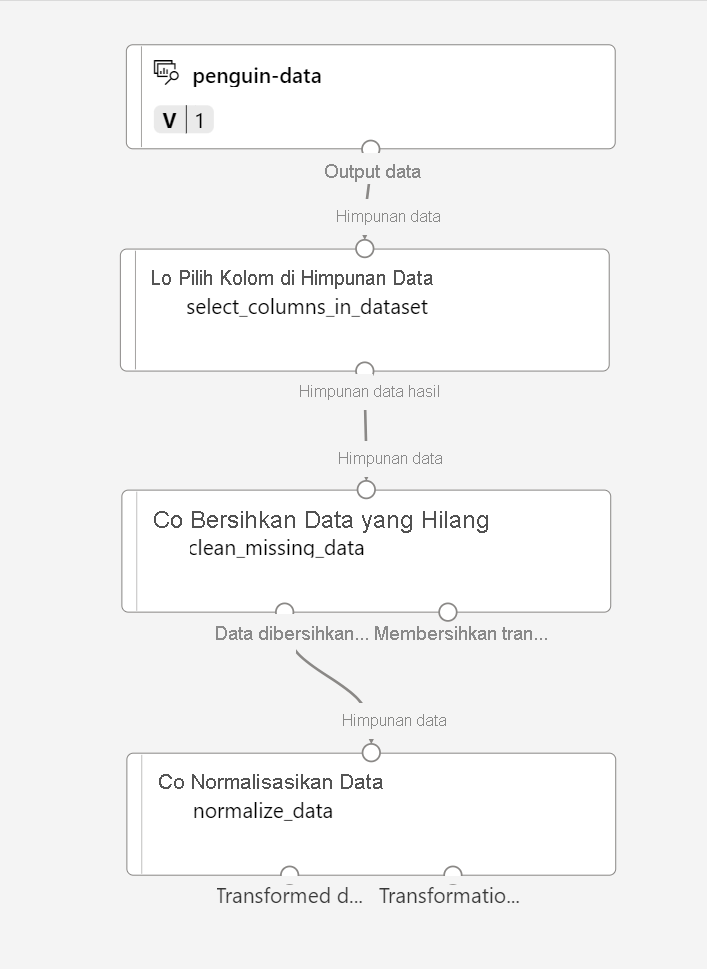
1. Klik dua kali modul **Membersihkan Data yang Hilang**, dan klik **Edit kolom** di panel pengaturan di sebelah kanan. Kemudian, di jendela **Pilih kolom,** pilih **Dengan aturan** dan sertakan **Semua kolom**; seperti ini:



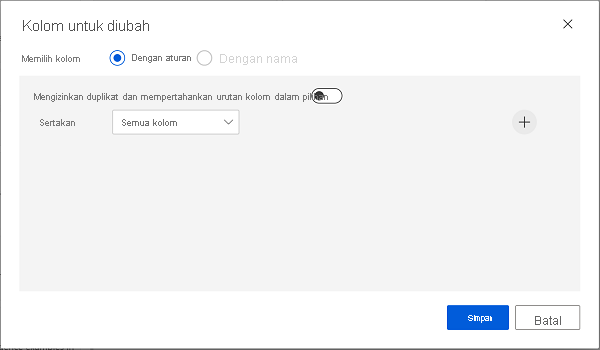
1. Dengan modul **Clean Missing Data** yang masih dipilih, di panel pengaturan, atur pengaturan konfigurasi berikut:

* **Rasio nilai hilang minimum:** 0,0
* **Rasio nilai maksimum yang hilang:** 1,0
* **Mode pembersihan**: Hapus seluruh baris

1. Di **Pustaka aset**, cari modul **Menormalisasi data** dan tempatkan ke kanvas, di bawah modul **Membersihkan Data yang Hilang**. Kemudian, sambungkan output paling kiri dari modul **Bersihkan Data yang Hilang** ke input modul **Normalisasi Data**.



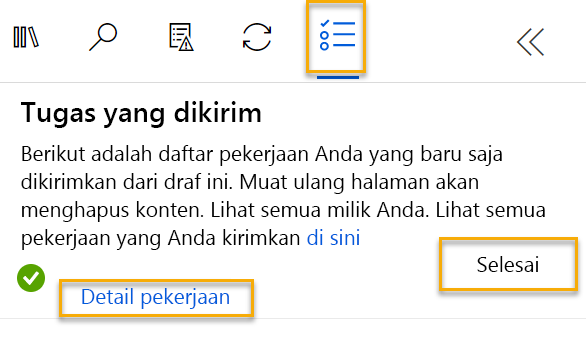
1. Klik dua kali modul **Menormalisasi Data**, dan di panel sebelah kanan, atur **Metode transformasi** ke **MinMax** dan pilih **Edit kolom.** Kemudian, di jendela **Pilih kolom**, pilih **Dengan aturan** dan sertakan **Semua kolom**; seperti ini:



1. Tutup pengaturan modul **Normalkan Data** untuk kembali ke kanvas perancang.
2. **Menjalankan alur**

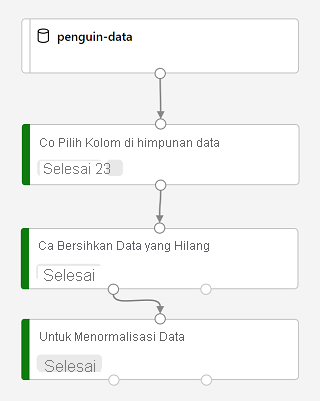
Untuk menerapkan transformasi data Anda, Anda perlu menjalankan alur sebagai eksperimen.

1. Pilih **Kirim**, dan jalankan alur sebagai **eksperimen baru** bernama **mslearn-penguin-training** di kluster komputasi Anda.
2. Tunggu sampai eksekusi selesai. Eksekusi eksperimen dapat memakan waktu 5 menit atau lebih.



Perhatikan bahwa panel sebelah kiri sekarang berada di panel **Pekerjaan yang Dikirim**. Anda akan mengetahui kapan eksekusi selesai karena status pekerjaan akan berubah menjadi **Selesai**.

1. **Menampilkan data yang diubah**
2. Setelah eksekusi selesai, himpunan data sekarang disiapkan untuk pelatihan model. Klik Detail Pekerjaan. Anda akan diarahkan ke jendela lain yang akan menampilkan modul seperti ini:



1. Di jendela baru, klik kanan modul **Menormalisasi Data**, pilih **Lakukan pratinjau data,** lalu pilih **Himpunan data yang ditransformasi** untuk melihat hasilnya.
2. Lihat data, mencatat bahwa kolom **Spesies** telah dihapus, tidak ada nilai yang hilang, dan nilai untuk keempat fitur telah dinormalisasi ke skala umum.
3. Menutup visualisasi hasil data yang dinormalisasi. Kembali ke jendela alur sebelumnya.

Setelah memilih dan menyiapkan fitur yang ingin Anda gunakan dari himpunan data, Anda siap menggunakannya untuk melatih model pengklusteran.

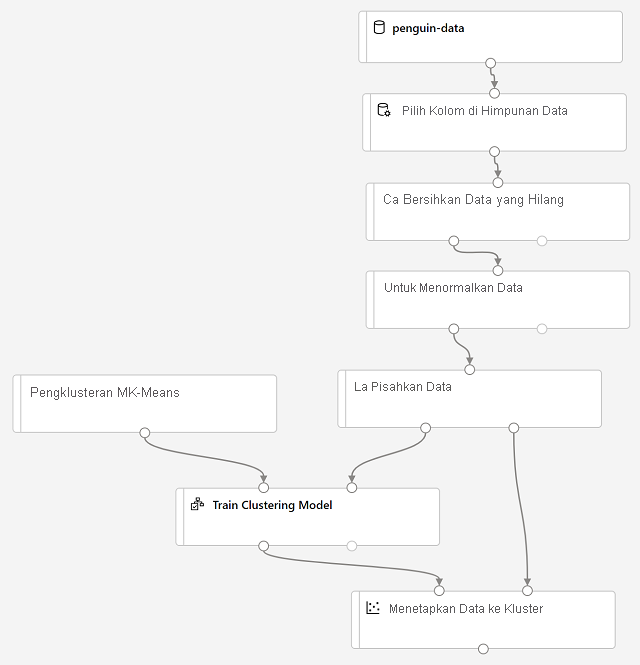
1. **Membuat dan menjalankan alur pelatihan**

Setelah menggunakan transformasi data untuk menyiapkan data, Anda dapat menggunakannya untuk melatih model pembelajaran mesin.

1. **Menambahkan modul pelatihan**

Untuk melatih model pengklusteran, Anda perlu menerapkan algoritme pengklusteran ke data, hanya menggunakan fitur yang telah Anda pilih untuk pengklusteran. Anda akan melatih model dengan subset data, dan menggunakan sisanya untuk menguji model terlatih.

Dalam latihan ini, Anda akan memperpanjang alur **Latih Pengklusteran Penguin** seperti yang ditunjukkan di sini:

****

Ikuti langkah-langkah di bawah, menggunakan gambar di atas untuk referensi saat Anda menambahkan dan mengonfigurasi modul yang diperlukan.

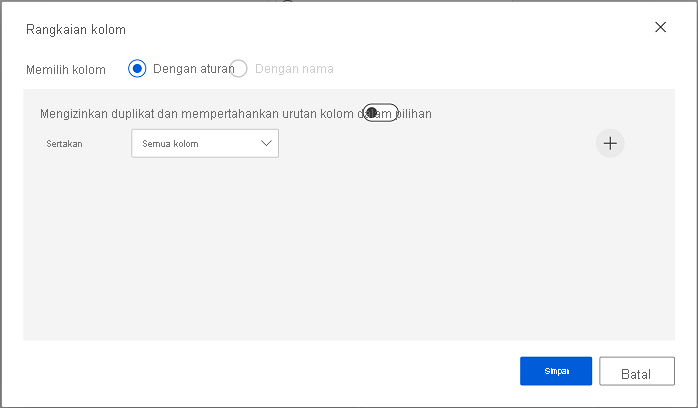
1. Buka alur Latih Pengklusteran Penguin, jika belum dibuka.
2. Di panel Pustaka Aset di sebelah kiri, cari dan tempatkan modul Memisahkan Data ke kanvas pada modul Menormalisasi Data. Kemudian, sambungkan output kiri modul Normalisasi Data ke input modul Pisahkan Data.

**Tip** : Gunakan bilah pencarian untuk menemukan modul dengan cepat.

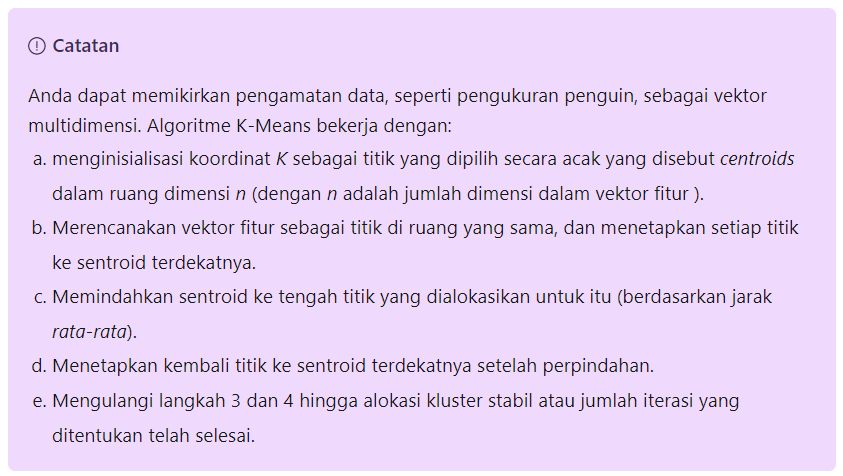
1. Pilih modul **Split Data**, dan konfigurasikan pengaturannya sebagai berikut:

* Mode pemisahan: Pisahkan Baris
* Pecahan baris dalam himpunan data output pertama: 0,7
* Pemisahan acak: True
* Nilai awal acak: 123
* Pemisahan bertingkat: False

1. Di **pustaka Aset,** cari dan tempatkan modul **Melatih Model Pengklusteran** ke kanvas, di bawah modul **Memisahkan Data**. Kemudian, sambungkan output Dataset1 hasil (kiri) modul **Pisahkan Data** ke input Himpunan data (kanan) modul **Latih Model Pengklusteran**.
2. Model pengklusteran harus menetapkan kluster ke item data dengan menggunakan semua fitur yang Anda pilih dari himpunan data asli. Klik dua kali modul Melatih Model Pengklusteran dan di panel kanan, pilih **Mengedit Kolom**. Gunakan opsi **Dengan aturan** untuk menyertakan semua kolom; seperti ini:



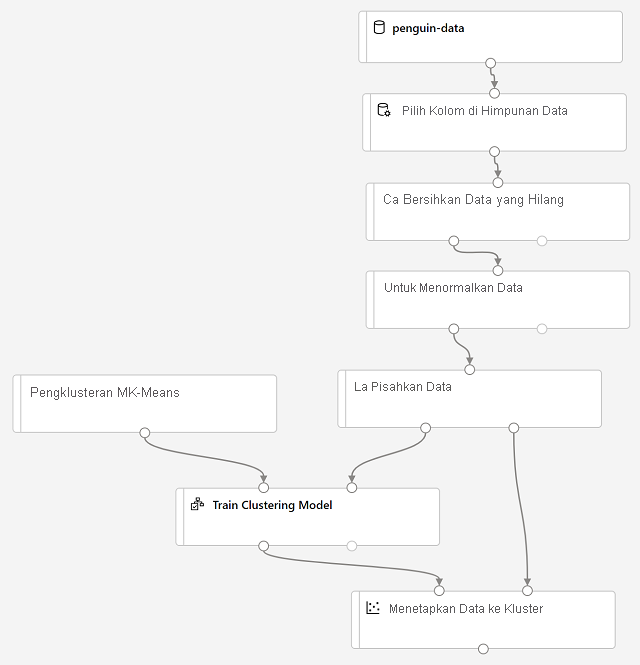
1. Model yang kami latih akan menggunakan fitur untuk mengelompokkan data ke dalam kluster, jadi kami perlu melatih model menggunakan algoritme pengklusteran. Dalam **Pustaka aset**, cari dan tempatkan modul **Pengklusteran K-Means ke kanvas**, ke sebelah kiri himpunan data **penguin-data** dan di atas modul **Melatih Model Pengklusteran**. Kemudian, sambungkan output-nya ke input **Model tidak terlatih (kiri**) modul **Latih Model Pengklusteran**.
2. Algoritme K-Means mengelompokkan item ke jumlah kluster yang Anda tentukan - nilai yang dirujuk sebagai **K**. Pilih modul **Pengklusteran K-Means** dan di kanan panel, atur parameter **Jumlah sentroid** menjadi **3.**



1. Setelah menggunakan 70% data untuk melatih model pengklusteran, Anda dapat menggunakan 30% sisanya untuk mengujinya menggunakan model untuk menetapkan data ke kluster. Dalam Pustaka aset, cari dan tempatkan modul Menetapkan Data ke Kluster ke kanvas, di bawah modul Melatih Model Pengklusteran. Kemudian, sambungkan output Model terlatih (kiri) modul Latih Model Pengklusteran ke input Model terlatih (kiri) modul Tetapkan Data ke Kluster; dan sambungkan output Dataset2 hasil (kanan) modul Pisahkan Data ke input Himpunan data (kanan) modul **Tetapkan Data ke Kluster.**
2. **Menjalankan alur pelatihan**

Sekarang Anda siap untuk menjalankan alur pelatihan dan melatih model.

1. Pastikan alur Anda seperti ini:



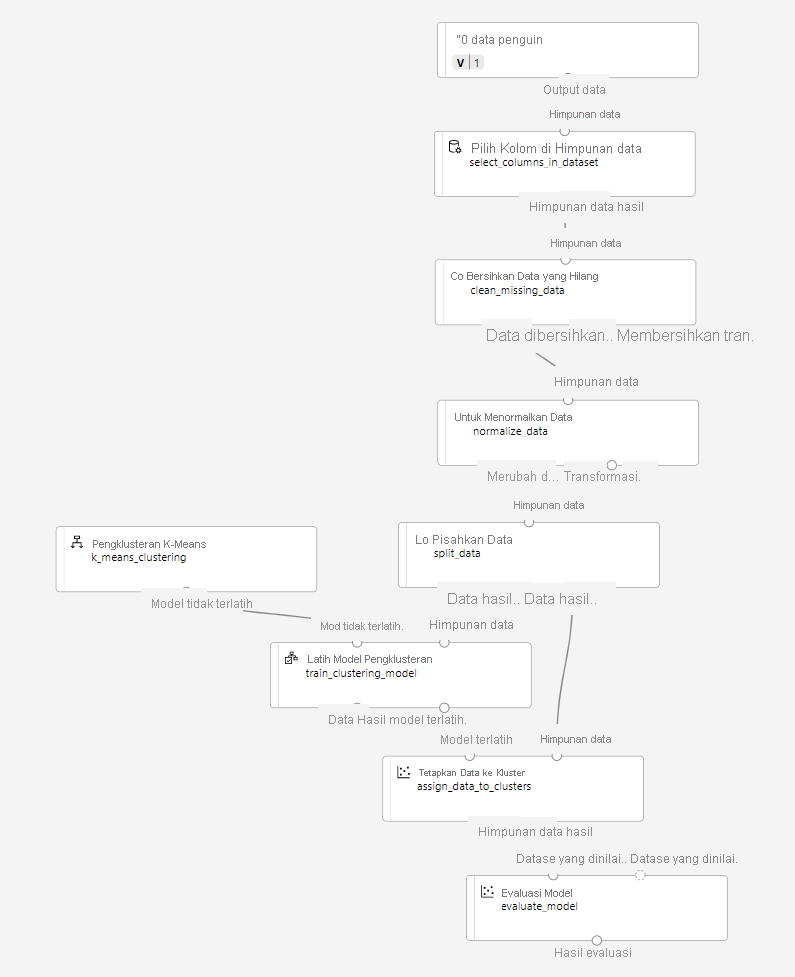
1. Pilih **Kirim**, dan jalankan alur menggunakan eksperimen yang ada bernama **mslearn-penguin-training** pada kluster komputasi Anda.
2. Tunggu hingga eksekusi eksperimen selesai. Eksekusi eksperimen dapat memakan waktu 5 menit atau lebih.
3. Setelah eksekusi eksperimen selesai, pilih Detail pekerjaan. Di jendela baru, klik kanan modul Menetapkan Data ke Kluster, pilih Pratinjau data, lalu pilih Himpunan data hasil untuk melihat hasilnya.
4. Gulir ke kanan, dan perhatikan kolom **Tugas**, yang berisi kluster (0, 1, atau 2) tempat setiap baris ditetapkan. Ada juga kolom baru yang menunjukkan jarak dari titik yang mewakili baris ini ke pusat masing-masing kluster - kluster yang titiknya paling dekat adalah yang ditetapkan.
5. Tutup visualisasi **Tetapkan Data ke Kluster.** Kembali ke jendela alur.

Model ini memprediksi kluster untuk pengamatan penguin, tetapi seberapa andal prediksinya? Untuk menilai itu, Anda perlu mengevaluasi model.

1. **Mengevaluasi model pengklusteran**

Mengevaluasi model pengklusteran dibuat sulit oleh fakta bahwa sebelumnya tidak ada nilai true yang diketahui untuk penetapan kluster. Model pengklusteran yang sukses adalah model yang mencapai tingkat pemisahan yang baik antara item di setiap kluster, jadi kami membutuhkan metrik untuk membantu kami mengukur pemisahan tersebut.

1. Buka alur **Latih Pengklusteran** Penguin yang Anda buat di unit sebelumnya jika belum dibuka.
2. Di Pustaka aset, cari dan tempatkan modul Mengevaluasi Model pada kanvas, di bawah modul Menetapkan Data ke Kluster. Sambungkan output modul Menetapkan Data ke Kluster ke input Himpunan data yang dinilai (kiri) dari modul Mengevaluasi Model.
3. Pastikan alur Anda terlihat seperti ini:



1. Pilih **Kirim**, dan jalankan alur menggunakan eksperimen **mslearn-penguin-training** yang ada.
2. Tunggu hingga eksekusi eksperimen selesai.
3. Setelah eksekusi eksperimen selesai, pilih Detail pekerjaan. Klik kanan modul Mengevaluasi Model dan pilih Pratinjau data, lalu pilih Hasil evaluasi. Metrik ini dapat membantu para ilmuwan data menilai seberapa baik model memisahkan kluster. Mereka menyertakan deretan metrik untuk setiap kluster, dan baris ringkasan untuk evaluasi gabungan. Metrik di setiap baris adalah:

* **Jarak Rata-Rata ke Pusat Lain**: Ini menunjukkan seberapa dekat, rata-rata, setiap titik dalam kluster dengan sentroid dari semua kluster lainnya.
* **Jarak Rata-rata ke Pusat Kluster**: Ini menunjukkan seberapa dekat, rata-rata, setiap titik dalam kluster dengan sentroid kluster.
* **Jumlah Titik**: Jumlah titik yang ditetapkan ke kluster.
* **Jarak Maksimal ke Pusat Kluster:** Maksimum jarak antara setiap titik dan sentroid dari kluster titik itu. Jika jumlah ini tinggi, kluster mungkin tersebar luas. Statistik ini yang dikombinasikan dengan **Jarak Rata-Rata ke Pusat Kluster** membantu Anda menentukan penyebaran kluster.

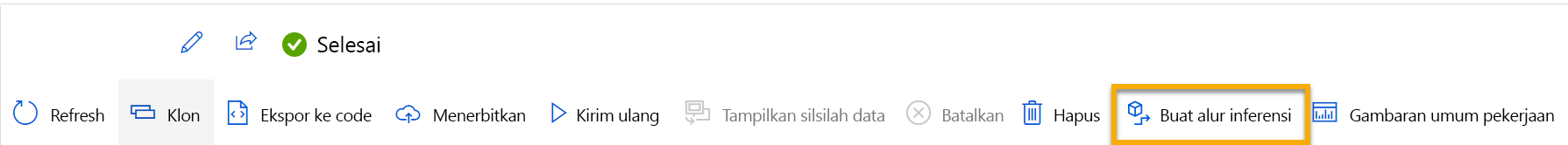
1. Tutup jendela visualisasi hasil Evaluasi Model.

Setelah memiliki model pengklusteran yang berfungsi, Anda dapat menggunakannya untuk menetapkan data baru ke kluster dalam alur inferensi.

1. **Membuat alur inferensi**

Setelah membuat dan menjalankan alur untuk melatih model pengklusteran, Anda dapat membuat alur inferensi. Alur inferensi menggunakan model untuk menetapkan pengamatan data baru ke kluster. Model ini akan menjadi dasar untuk layanan prediktif yang dapat Anda terbitkan untuk digunakan oleh aplikasi.

1. **Membuat alur inferensi**
2. Di Azure Machine Learning studio, luaskan panel sebelah kiri dengan memilih tiga baris di kiri atas layar. Klik Pekerjaan (di bagian Aset) untuk melihat semua pekerjaan yang telah Anda jalankan. Pilih eksperimen mslearn-penguin-training, lalu pilih alurmslearn-penguin-training.
3. Cari menu di atas kanvas dan klik Buat alur inferensi. Anda mungkin perlu memperluas layar hingga penuh dan mengeklik ikon tiga titik ... di sudut kanan atas layar untuk menemukan Buat alur inferensi di menu.

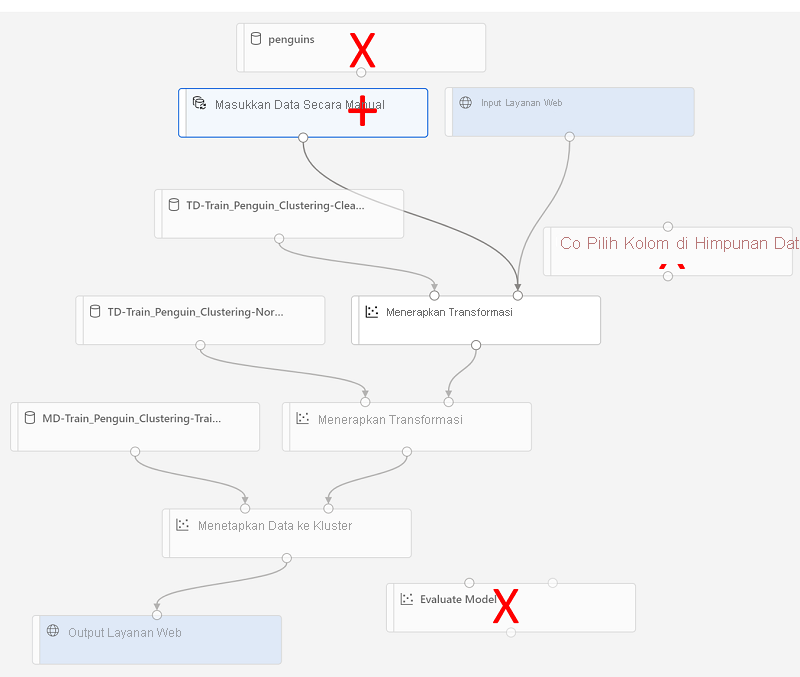


1. Di daftar drop-down Buat alur inferensi, klik Alur inferensi real time. Setelah beberapa detik, versi baru dari alur Anda bernama Latih Pengklusteran Penguin-inferensi real time akan dibuka.

**Jika alur tidak menyertakan modul Input Layanan Web dan Output Layanan Web**, kembali ke halaman Perancang lalu buka kembali Latih alur inferensi real time Pengklusteran Penguin.

1. Buka Pengaturan di menu kanan atas. Pada Detail draft, ganti nama alur baru menjadi Prediksi Kluster Penguin, lalu tinjau alur baru. Ini berisi input layanan web untuk data baru yang akan dikirimkan, dan output layanan web untuk menampilkan hasil. Model transformasi dan pengklusteran dalam alur pelatihan Anda adalah bagian dari alur ini. Model yang dilatih akan digunakan untuk menilai data baru.

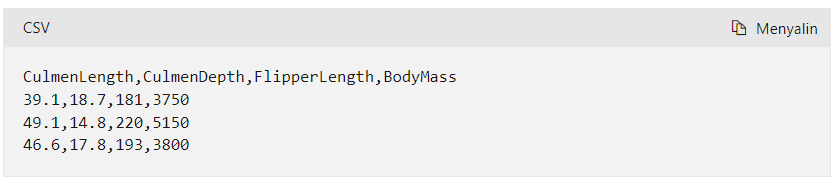
Anda akan membuat perubahan berikut pada alur inferensi:



* Ganti himpunan data penguin-data dengan modul Masukkan Data Secara Manual yang tidak menyertakan kolom Spesies.
* Hapus modul Pilih Kolom dalam Himpunan Data, yang sekarang redundan.
* Sambungkan modul Input Layanan Web dan Masukkan Data Secara Manual (yang mewakili input data yang akan diklusterkan) ke modul Terapkan Transformasi pertama.
* Hapus modul Evaluasi Model.

Ikuti langkah-langkah yang tersisa di bawah, menggunakan gambar dan informasi di atas untuk referensi saat Anda memodifikasi alur.

1. Alur inferensi mengasumsikan bahwa data baru akan cocok dengan skema data pelatihan asli, sehingga himpunan data penguin-data dari alur pelatihan disertakan. Namun, data input ini mencakup kolom untuk spesies penguin, yang tidak digunakan model. Hapus himpunan data penguin-data dan modul Pilih Kolom dalam Himpunan Data, dan ganti dengan modul Masukkan Data Secara Manual dari bagian Pustaka aset. Kemudian, ubah pengaturan modul Masukkan Data Secara Manual untuk menggunakan input CSV berikut, yang berisi nilai fitur untuk tiga pengamatan penguin baru (termasuk header):



1. Sambungkan output dari modul Input Layanan Web dan Masukkan Data Secara Manual ke input Himpunan Data (kanan) modul Terapkan Transformasi pertama.
2. Hapus modul Evaluasi Model.
3. Verifikasi bahwa alur Anda terlihat mirip dengan gambar berikut:
4. Kirimkan alur sebagai eksperimen baru bernama mslearn-penguin-inference pada kluster komputasi Anda. Eksperimen mungkin memerlukan beberapa saat untuk dijalankan.
5. Setelah alur selesai, pilih Detail pekerjaan. Di jendela baru, klik kanan modul Menetapkan Data ke Kluster dan pilih Lakukan pratinjau data untuk melihat penetapan dan metrik kluster yang diprediksi untuk tiga pengamatan penguin dalam data input.

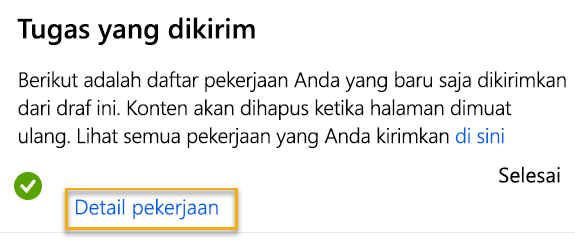
Alur inferensi Anda menetapkan pengamatan penguin ke kluster berdasarkan fiturnya. Sekarang Anda siap untuk menerbitkan alur sehingga aplikasi klien dapat menggunakannya.

1. **Menyebarkan layanan prediktif**

Setelah membuat dan menguji alur inferensi untuk inferensi real time, Anda bisa menerbitkannya sebagai layanan untuk digunakan aplikasi klien.

**Catatan**: Dalam latihan ini, Anda akan menyebarkan layanan web ke Azure Container Instance (ACI). Jenis komputasi ini dibuat secara dinamis, dan berguna untuk pengembangan dan pengujian. Untuk produksi, Anda harus membuat kluster inferensi, yang menyediakan kluster Azure Kubernetes Service (AKS) yang memberikan skalabilitas dan keamanan yang lebih baik.

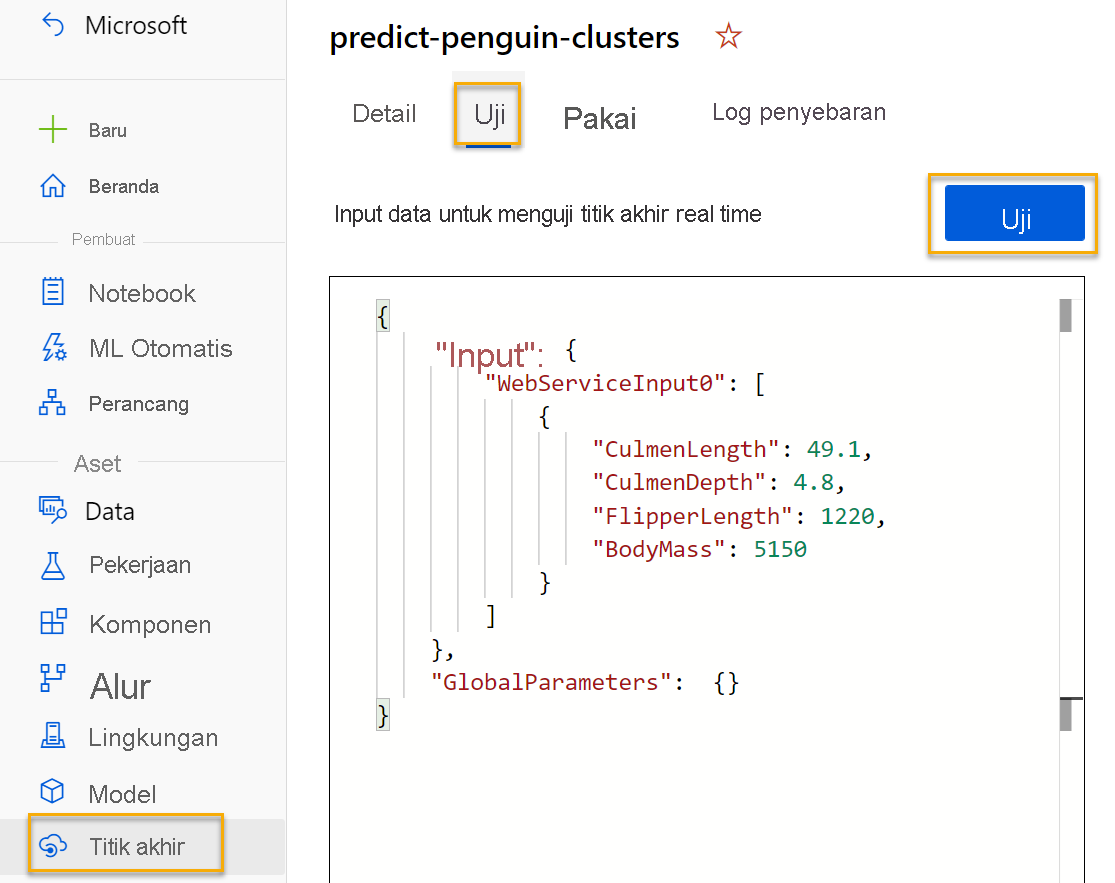
1. **Menyebarkan layanan**
2. Lihat alur inferensi **Prediksi Kluster Penguin** yang Anda buat di unit sebelumnya.
3. Pilih **Detail pekerjaan** di panel sebelah kiri. Panel akan membuka jendela lain.



1. Di jendela baru, pilih Sebarkan.
2. Di kanan atas, pilih **Sebarkan**, dan sebarkan titik akhir real time baru, menggunakan pengaturan berikut:

* **Nama**: predict-penguin-clusters
* **Deskripsi**: Penguin kluster.
* **Jenis komputasi**: Azure Container Instance

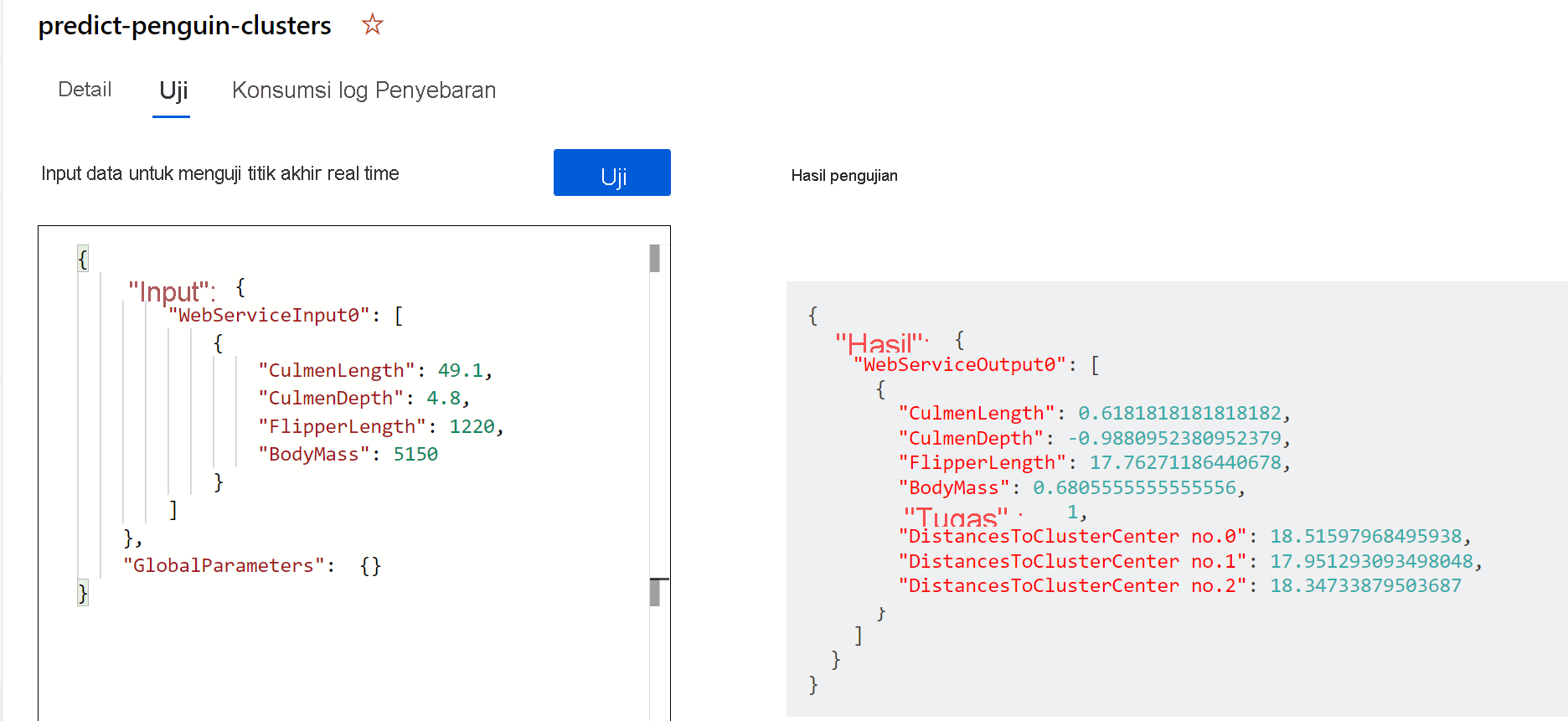
1. Tunggu hingga layanan web disebarkan - ini bisa memakan waktu beberapa menit. Status penyebaran ditampilkan di kiri atas antarmuka perancang.
2. **Menguji layanan**
3. Di halaman Titik akhir, buka titik akhir real time **predict-penguin-clusters.**



1. Saat titik akhir predict-auto-price terbuka, pilih tab Uji. Kita akan menggunakannya untuk menguji model dengan data baru. Hapus data saat ini pada Data input untuk menguji titik akhir real time. Salin dan tempel data di bawah ini ke bagian data:



1. Pilih **Uji**. Di sisi kanan layar, Anda akan melihat 'penetapan' output. Perhatikan bagaimana kluster yang ditetapkan adalah kluster dengan jarak terpendek ke pusat kluster.



Anda baru saja menguji layanan yang siap untuk disambungkan ke aplikasi klien menggunakan kredensial di tab Konsumsi. Kami akan mengakhiri lab di sini. Anda boleh terus bereksperimen dengan layanan yang baru saja disebarkan.

1. **Ringkasan**

Dalam modul ini, Anda mempelajari cara menggunakan perancang Azure Machine Learning untuk melatih dan menerbitkan model pengklusteran.

1. **Pembersihan**

Layanan web yang Anda buat dihosting dalam Azure Container Instance. Jika tidak berniat untuk bereksperimen dengan ini lebih lanjut, Anda harus menghapus titik akhir untuk menghindari mengumpulkan penggunaan Azure yang tidak perlu. Anda juga harus menghentikan instans komputasi hingga membutuhkannya lagi.

1. Di studio Azure Machine Learning, di tab Titik Akhir, pilih titik akhir predict-penguin-clusters. Kemudian pilih **Hapus** () dan konfirmasikan bahwa Anda ingin menghapus titik akhir.
2. Di halaman Komputasi, di tab Instans Komputasi, pilih instans komputasi Anda lalu pilih Hentikan.

