Menjelaskan model pembelajaran mesin dengan Azure Machine Learning

Banyak keputusan yang dibuat oleh organisasi dan sistem otomatis saat ini didasarkan pada prediksi yang dibuat oleh model pembelajaran mesin. Semakin penting untuk dapat memahami faktor-faktor yang mempengaruhi model prediksi yang dibuat.

**Tujuan pembelajaran**

Pelajari cara menjelaskan model dengan menghitung dan menginterpretasikan pentingnya fitur.

1. **Pengantar**

Karena pembelajaran mesin menjadi semakin integral dengan keputusan yang memengaruhi kesehatan, keselamatan, kesejahteraan ekonomi, dan aspek lain dari kehidupan masyarakat, penting untuk dapat memahami bagaimana model membuat prediksi; dan untuk dapat menjelaskan alasan keputusan berbasis machine learning.

Menjelaskan model sulit karena berbagai jenis algoritme pembelajaran mesin dan sifat cara kerja pembelajaran mesin, tetapi interpretasi model telah menjadi elemen kunci untuk membantu membuat prediksi model dapat dijelaskan.

1. **Tujuan pembelajaran**

Dalam modul ini, Anda akan mempelajari cara:

* Menginterpretasikan kepentingan fitur global dan lokal.
* Gunakan penjelajah untuk menginterpretasikan model.
* Buat penjelasan model dalam eksperimen pelatihan.
* Visualisasikan penjelasan model.

1. **Kepentingan fitur**

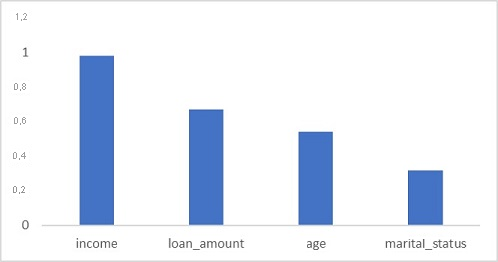
Penjelas model menggunakan teknik statistik untuk menghitung kepentingan fitur. Hal ini memungkinkan Anda untuk mengukur pengaruh relatif setiap fitur dalam himpunan data pelatihan terhadap prediksi label. Penjelas bekerja dengan mengevaluasi himpunan data uji kasus fitur dan label yang diprediksi model untuknya.

1. **Kepentingan fitur global**

Kepentingan fitur global mengukur kepentingan relatif dari setiap fitur dalam himpunan data pengujian secara keseluruhan. Hal ini memberikan perbandingan umum sejauh mana setiap fitur dalam himpunan data memengaruhi prediksi.

Misalnya, model klasifikasi biner untuk memprediksi risiko gagal bayar pinjaman mungkin dilatih dari fitur seperti jumlah pinjaman, pendapatan, status perkawinan, dan usia untuk memprediksi label 1 untuk pinjaman yang kemungkinan akan dilunasi, dan 0 untuk pinjaman yang memiliki risiko default yang signifikan (dan karenanya tidak boleh disetujui). Penjelas kemudian dapat menggunakan himpunan data pengujian yang cukup representatif untuk menghasilkan nilai kepentingan fitur global berikut:

* pendapatan: 0,98
* jumlah pinjaman: 0,67
* usia: 0,54
* status perkawinan 0,32

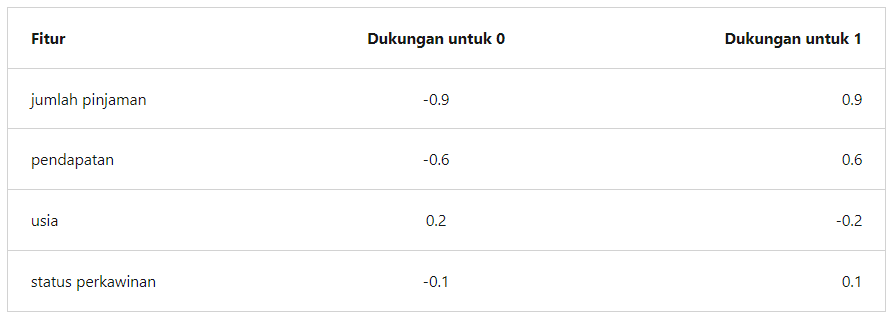


Jelas dari nilai-nilai ini, bahwa sehubungan dengan prediksi keseluruhan yang dihasilkan oleh model untuk himpunan data pengujian, **pendapatan** adalah fitur paling penting untuk memprediksi apakah peminjam akan gagal membayar pinjaman atau tidak, diikuti oleh **jumlah pinjaman**, lalu **usia**, dan terakhir status **perkawinan**.

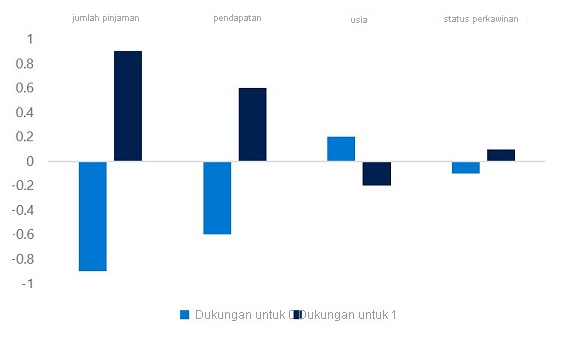
1. **Kepentingan fitur lokal**

Kepentingan fitur lokal mengukur pengaruh setiap nilai fitur untuk prediksi individu tertentu.

Misalnya, Sam mengajukan pinjaman, yang disetujui oleh model pembelajaran mesin (dengan memprediksi bahwa Sam tidak akan gagal membayar pinjaman). Anda dapat menggunakan penjelas untuk menghitung kepentingan fitur lokal untuk aplikasi Sam untuk menentukan faktor mana yang mempengaruhi prediksi (dengan memprediksi bahwa Sam tidak akan gagal membayar pinjaman). Anda mungkin mendapatkan hasil seperti ini:



Karena ini adalah model klasifikasi, setiap fitur mendapatkan nilai kepentingan lokal untuk setiap kelas yang memungkinkan, yang menunjukkan jumlah dukungan untuk kelas tersebut berdasarkan nilai fitur. Karena ini adalah model klasifikasi biner, hanya ada dua kemungkinan kelas (0 dan 1). Dukungan setiap fitur untuk satu kelas menghasilkan tingkat dukungan yang relatif negatif untuk kelas lainnya.



Dalam kasus Sam, dukungan keseluruhan untuk kelas 0 adalah -1.4, dan dukungan untuk kelas 1 adalah 1.4; jadi dukungan untuk kelas 1 lebih tinggi daripada untuk kelas 0, dan pinjaman disetujui. Fitur terpenting untuk prediksi kelas 1 adalah **jumlah pinjaman,** diikuti oleh **pendapatan** - ini adalah urutan yang berlawanan dari nilai kepentingan fitur globalnya (yang menunjukkan bahwa **pendapatan** adalah faktor terpenting untuk sampel data secara keseluruhan). Mungkin ada beberapa alasan mengapa kepentingan lokal untuk prediksi individu bervariasi dari kepentingan global untuk keseluruhan himpunan data; misalnya, Sam mungkin memiliki pendapatan yang lebih rendah daripada rata-rata, tetapi jumlah pinjaman dalam kasus ini mungkin sangat kecil.

Untuk model klasifikasi multi-kelas, nilai kepentingan lokal untuk setiap kelas yang mungkin dihitung untuk setiap fitur, dengan total di semua kelas selalu 0. Misalnya, model dapat memprediksi spesies penguin berdasarkan fitur seperti panjang paruh, lebar paruh, panjang sirip, dan berat. Misalkan ada tiga spesies penguin, maka model memprediksi salah satu dari tiga label kelas (0, 1, atau 2). Untuk prediksi individu, fitur panjang sirip mungkin memiliki nilai kepentingan lokal 0,5 untuk kelas 0, 0,3 untuk kelas 1, dan -0,8 untuk kelas 2 - menunjukkan bahwa panjang sirip cukup mendukung prediksi kelas 0, sedikit mendukung prediksi kelas 1, dan sangat mendukung prediksi bahwa penguin ini **bukan** kelas 2.

Untuk model regresi, tidak ada kelas sehingga nilai kepentingan lokal hanya menunjukkan tingkat pengaruh setiap fitur pada label skalar yang diprediksi.

1. **Menggunakan penjelas**

Anda dapat menggunakan SDK Azure Machine Learning untuk membuat penjelas bagi model, meskipun mereka tidak dilatih menggunakan eksperimen Azure Machine Learning.

1. **Membuat penjelas**

Untuk menginterpretasikan model lokal, Anda harus menginstal paket **azureml-interpret** dan menggunakannya untuk membuat penjelas. Ada beberapa jenis penjelas, antara lain:

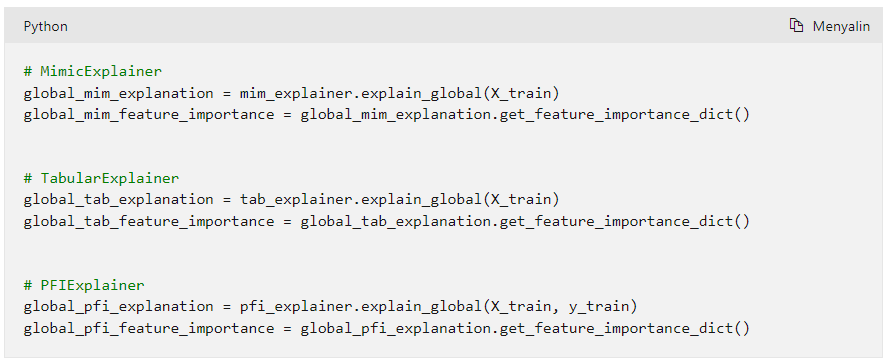
* **MimicExplainer** - Penjelas yang membuat model pengganti global yang mendekati model terlatih Anda dan dapat digunakan untuk menghasilkan penjelasan. Model yang dapat dijelaskan ini harus memiliki jenis arsitektur yang sama dengan model terlatih Anda (misalnya, linier atau berbasis pohon).
* **TabularExplainer** - Penjelas yang bertindak sebagai pembungkus berbagai algoritme penjelas SHAP, secara otomatis memilih salah satu yang paling sesuai untuk arsitektur model Anda.
* **PFIExplainer** - penjelas Kepentingan Fitur Permutasi yang menganalisis kepentingan fitur dengan mengacak nilai fitur dan mengukur dampaknya pada performa prediksi.

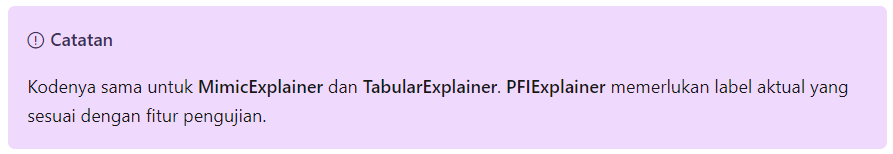
Contoh kode berikut menunjukkan cara membuat instans dari masing-masing jenis penjelas ini untuk model hipotetis bernama **loan\_model**:



1. **Menjelaskan kepentingan fitur global**

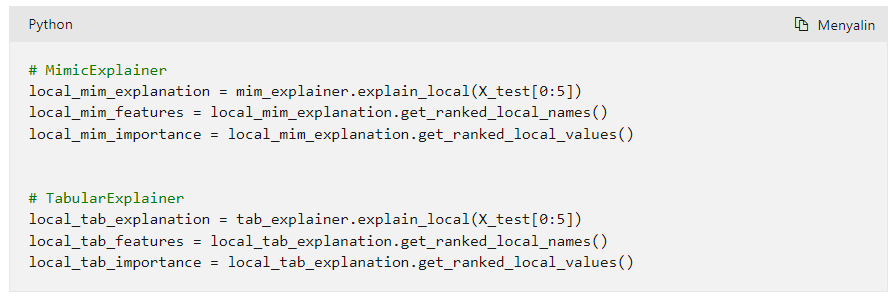
Untuk mengambil nilai kepentingan global untuk fitur dalam model Anda, panggil metode **explain\_global()** dari penjelas Anda untuk mendapatkan penjelasan global, lalu gunakan metode **get\_feature\_importance\_dict()** untuk mendapatkan kamus nilai-nilai kepentingan fitur. Contoh kode berikut menunjukkan cara mengambil kepentingan fitur global:

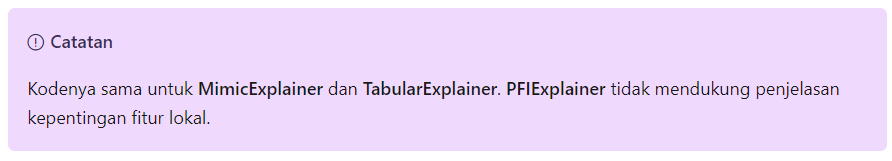




1. **Menjelaskan kepentingan fitur lokal**

Untuk mengambil kepentingan fitur lokal dari **MimicExplainer** atau **TabularExplainer**, Anda harus memanggil metode **explain\_local()** dari penjelas Anda, dengan menetapkan subset kasus yang ingin Anda jelaskan. Kemudian Anda dapat menggunakan metode **get\_ranked\_local\_names()** dan **get\_ranked\_local\_values()** untuk mengambil kamus nama fitur dan nilai penting, yang diberi peringkat berdasarkan kepentingan. Contoh kode berikut menunjukkan cara mengambil kepentingan fitur lokal:



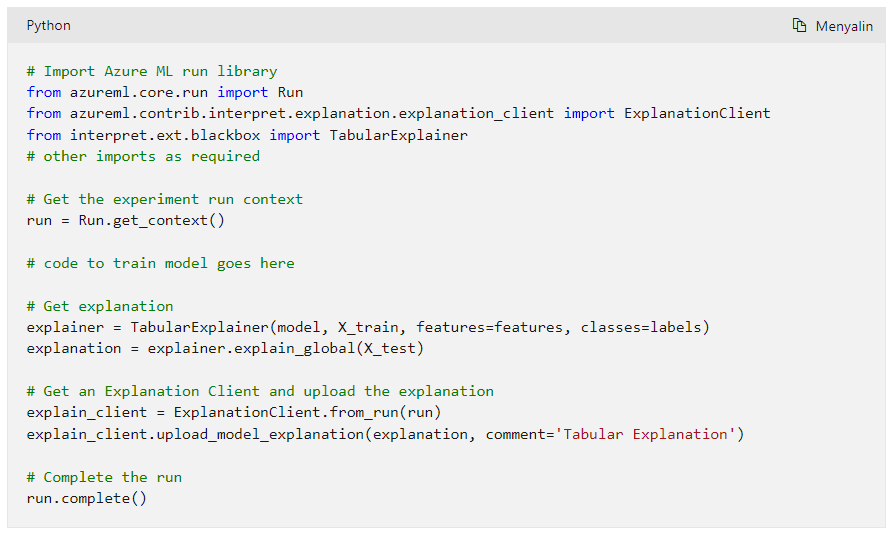


1. **Membuat penjelasan**

Saat Anda menggunakan penghitung atau skrip untuk melatih model dalam eksperimen Azure Machine Learning, Anda bisa membuat penjelas dan mengunggah penjelasan yang dihasilkannya ke proses untuk analisis nanti.

1. **Membuat penjelasan dalam skrip eksperimen**

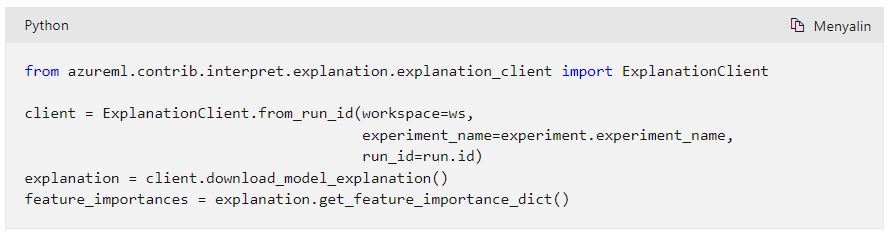
Untuk membuat penjelasan dalam skrip eksperimen, Anda harus memastikan bahwa paket **azureml-interpret** dan **azureml-contrib-interpret** diinstal di lingkungan yang dijalankan. Kemudian Anda dapat menggunakan ini untuk membuat penjelasan dari model terlatih Anda dan mengunggahnya ke output yang dijalankan. Contoh kode berikut menunjukkan bagaimana kode untuk membuat dan mengunggah penjelasan model dapat dimasukkan ke dalam skrip eksperimen.



1. **Melihat penjelasan**

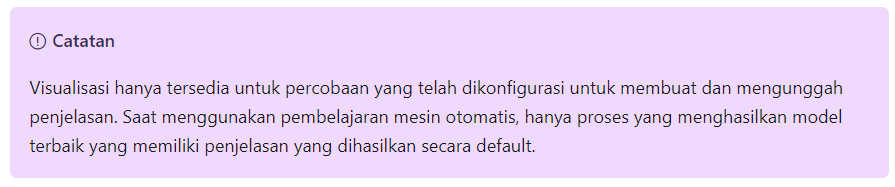
Anda dapat melihat penjelasan yang Anda buat untuk model Anda di tab **Penjelasan** untuk dijalankan di studio Azure Machine Learning.

Anda juga dapat menggunakan objek **ExplanationClient** untuk mengunduh penjelasan dengan Python.



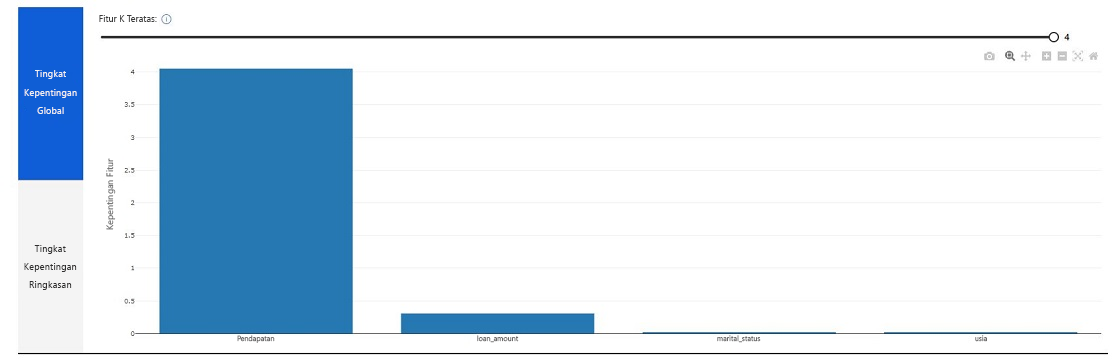
1. **Memvisualisasikan penjelasan**

Penjelasan model di studio Azure Machine Learning menyertakan beberapa visualisasi yang dapat Anda gunakan untuk menjelajahi kepentingan fitur.



1. **Memvisualisasikan kepentingan fitur global**

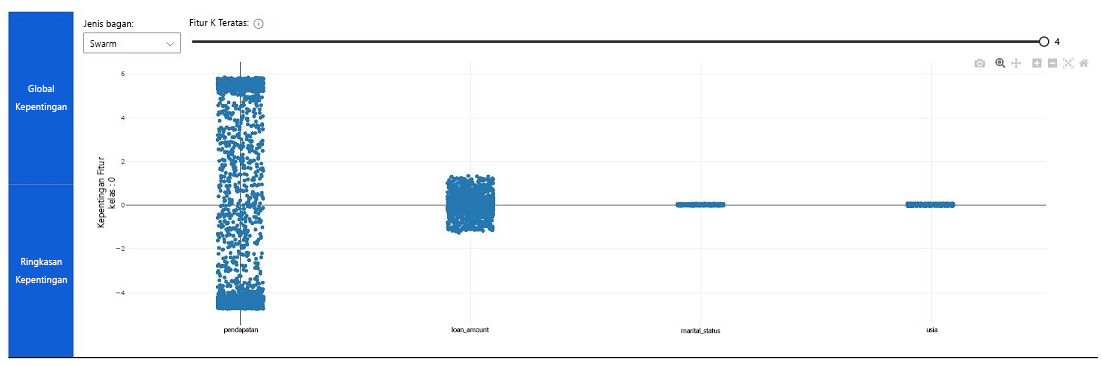
Visualisasi pertama pada tab **Penjelasan** untuk proses menunjukkan kepentingan fitur global.



Anda dapat menggunakan penggeser untuk hanya menampilkan fitur N teratas.

1. **Memvisualisasikan ringkasan kepentingan**

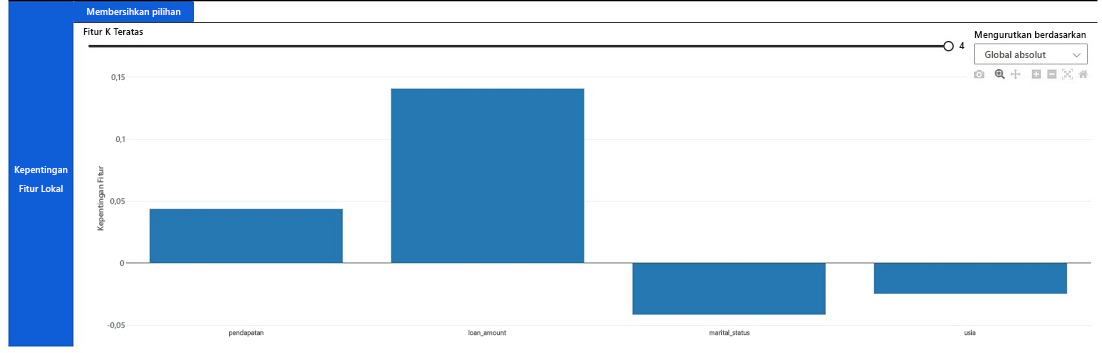
Beralih ke visualisasi **Ringkasan Kepentingan** menunjukkan distribusi nilai kepentingan individu untuk setiap fitur di seluruh himpunan data pengujian.



Anda dapat melihat fitur sebagai plot swarm (ditampilkan di atas), plot kotak, atau plot biola.

1. **Memvisualisasikan kepentingan fitur lokal**

Memilih titik data individual menunjukkan kepentingan fitur lokal untuk kasus tempat titik data tersebut berada.



1. **Latihan - Menginterpretasikan model**

Sekarang adalah kesempatan Anda untuk menginterpretasikan model.

Dalam latihan ini, Anda akan:

* Hasilkan kepentingan fitur untuk model.
* Hasilkan penjelasan sebagai bagian dari eksperimen pelatihan model.

1. **Instruksi**

Ikuti petunjuk ini untuk menyelesaikan latihan.

1. Jika Anda belum memiliki langganan Azure, daftar untuk uji coba gratis di <https://azure.microsoft.com.>
2. Menampilkan repo latihan di <https://aka.ms/mslearn-dp100.> \ <https://microsoftlearning.github.io/mslearn-dp100/>
3. Jika Anda belum melakukannya, selesaikan latihan **Membuat ruang kerja Azure Machine Learning** untuk memprovisikan ruang kerja Azure Machine Learning, membuat instans komputasi, dan mengkloning file yang diperlukan.
4. Selesaikan latihan **Menginterpretasikan model.**
5. **Ringkasan**

Dalam modul ini, Anda belajar cara:

* Menginterpretasikan kepentingan fitur global dan lokal.
* Gunakan penjelajah untuk menginterpretasikan model.
* Buat penjelasan model dalam eksperimen pelatihan.
* Visualisasikan penjelasan model.

Untuk mempelajari selengkapnya tentang menginterpretasikan model, lihat **Interpretasi model dalam Azure Machine Learning** dalam dokumentasi Azure Machine Learning.

Link : <https://learn.microsoft.com/id-id/azure/machine-learning/how-to-machine-learning-interpretability>