

Nama: Kevin Avicenna Widiarto

Portofolio

kevinavicenna.github.io

<https://www.linkedin.com/in/kevinavicenna/>

<https://github.com/kevinavicenna/>

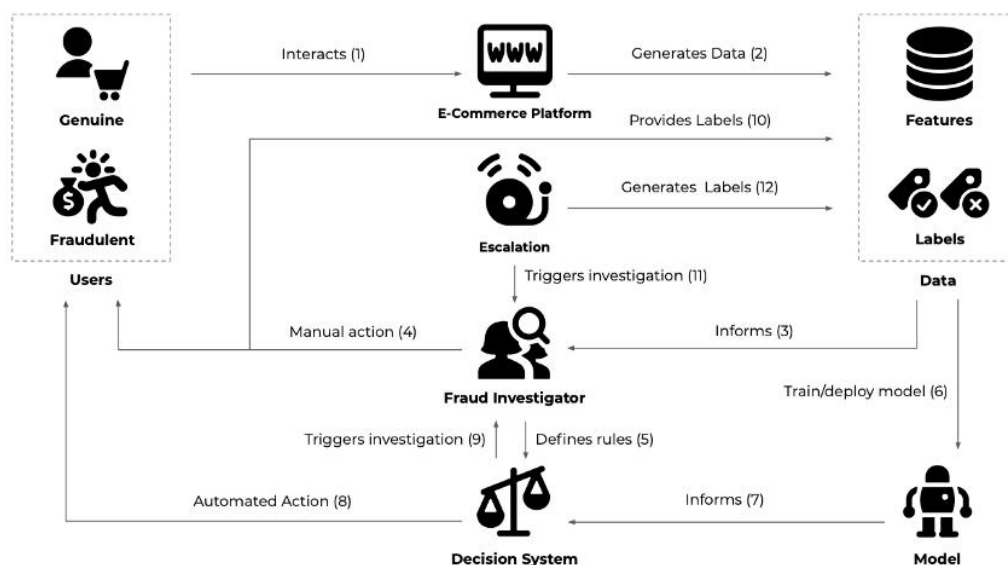
Use Case Retail Fraud Detection

Deteksi Penipuan dalam Retail Menggunakan KNN

Data Science di retail membantu melindungi reputasi perusahaan . Deteksi penipuan menjadi masalah yang sangat menantang bagi retail yang akibatnya mengalami kerugian finansial besar perusahaan ,lalu perusahaan retail mengambil bantuan dari teknologi digital baru yaitu konsep pembelajaran mesin atau Mechine Learning dan jaringan saraf atau Neural Network. Ini memungkinkan mereka untuk selalu memeriksa semua aktivitas dan menangkap aktivitas Penipuan. Penipuan juga diisolasi menggunakan beberapa teknik deteksi outlier. Alat deteksi outlier memiliki cara sendiri untuk mengatasi masalah, seperti analisis deret waktu(Time Series), analisis klaster(Claster Analyst), pemantauan transaksi waktu nyata, dll. Dan ketika kita berbicara tentang teknik pembelajaran mesin untuk memecahkan masalah ini.

Setelah hari demi hari penipuan semakin marak terjadi, banyak penipuan terjadi disebabkan oleh transaksi online. Situs Cekrekening.id mengumumkan kasus penipuan online dari ecommerce dan jualan online di media sosial dilaporkan hingga September 2021 sebanyak 115.756 kasus. Kasus yang dilaporkan tersebut terjadi penurunan jika dibandingkan dengan tahun 2020 yang mencapai 160 ribu lebih kasus. Modus tersebut dilakukan karena e-commerce baru akan memberikan uang dari pembeli ke penjual setelah transaksi berhasil, sehingga tidak memungkinkan untuk melakukan penipuan.

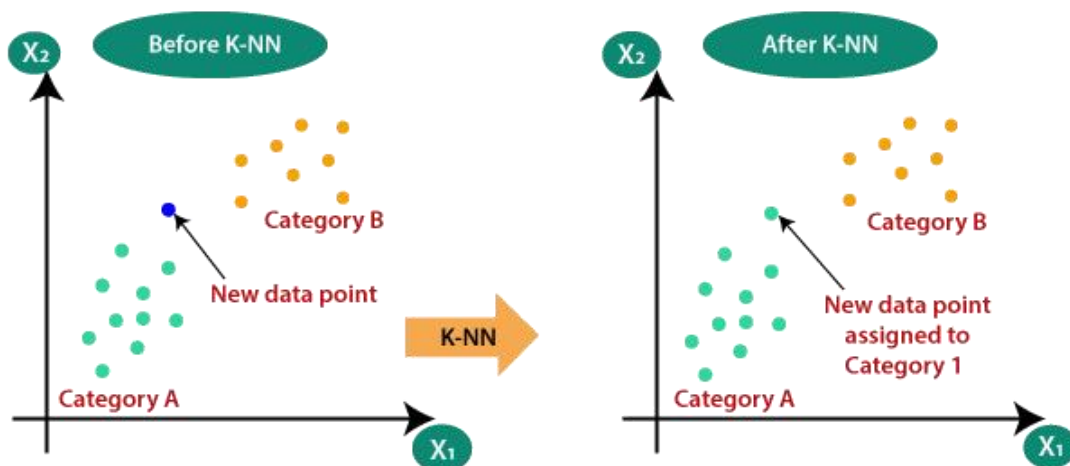
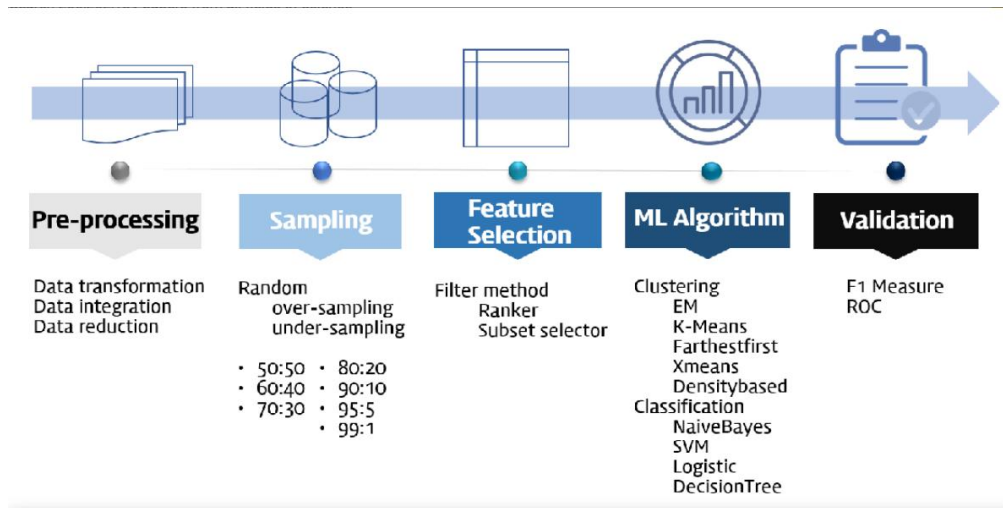
Selain itu, modus lain yang dilakukan pelaku kejahatan adalah mengambil alih akun e-commerce dan mengambil saldo uang digital yang ada di akun, Namun jumlah kasus ini relatif sedikit.



Maka dari itu menggunakan teknologi Machine Learning dan Neural Network sangat penting bagi retail, dengan ada banyak nya penipuan dapat digunakan pengaplikasian algoritma Machine learning salah satunya KNN atau yang disebut Kmean Clustering merupakan jenis Unsupervised Learning, yang digunakan ketika Anda memiliki data yang tidak berlabel. Hasil dari algoritma clustering K-means adalah :

- Centroid dari K cluster, digunakan untuk memberi label pada data baru
- Label untuk data training (setiap titik data ditugaskan ke satu cluster)

Manfaat KNN dapat melakukan pengelompokan untuk pengecekan apakah ada nya terdapat nilai outlier.



K-Nearest Neighbor

Berikut teknik cara penggunaan algoritma [KNN untuk mendeteksi fraud](#)

1. Melakukan data preprocessing
2. Mengimport library yang diperlukan seperti pandas, numpy, time matplotlib dan sklearn
3. Membaca data, apakah ada mencari missing value
4. Melabeli sebuah data
5. Ensembling model
6. Setting express verification

Dengan ada nya algoritma ini bisa dapat ditarik kesimpulan untuk mendeteksi nilai outlier , untuk kelebihan dan kekurangan dari algoritma ini adalah diantara nya :

Kelebihan: Metode ini tidak sensitif terhadap data yang hilang dan berisik, yang memungkinkan untuk mengonfigurasi kumpulan data yang lebih besar dengan persiapan yang lebih sedikit. Ini juga dianggap sangat akurat dan tidak memerlukan banyak upaya rekayasa untuk mengubah model.

Kekurangan: Seperti jaringan saraf, tetangga k-terdekat membutuhkan infrastruktur yang kuat dan mereka juga tidak memiliki interpretasi.

Kesimpulan nya Data Science di retail membantu melindungi reputasi perusahaan. Penipuan juga diisolasi menggunakan beberapa teknik deteksi outlier. Modus tersebut dilakukan karena e-commerce baru akan memberikan uang dari pembeli ke penjual. Dengan salah satunya dengan menggunakan KNN dapat mendeteksi nilai outlier sehingga meminimalisir terjadinya fraud