

## DAFTAR SINGKATAN

CNN	:	Convolutional Neural Network
EEG	:	Electroencephalogram
Hz	:	Hertz
WHO	:	World Health Organization
MEG	:	Magneto-Encephalography
fMRI	:	Functional Magnetic Resonance Imaging
1D	:	1 Dimensi

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Epilepsi adalah gangguan neurologis kronis yang menandai dirinya dengan serangan atau kejang berulang yang disebabkan oleh aktivitas listrik abnormal di otak. Gangguan ini dapat mempengaruhi individu dari berbagai kelompok usia dan latar belakang. Kejang epilepsi dapat bervariasi dalam tingkat keparahan, mulai dari kejang kecil yang hanya memengaruhi bagian kecil tubuh hingga kejang besar yang melibatkan seluruh tubuh. Salah satu aspek penting dalam memahami epilepsi adalah menyadari efek dan risikonya.

Efek epilepsi dapat bervariasi, termasuk gangguan fungsi otak seperti kesulitan konsentrasi, masalah belajar, dan gangguan memori. Selama kejang, penderita epilepsi juga dapat mengalami cedera fisik seperti jatuh atau terluka akibat kejang yang tidak terkendali. Selain dampak fisik, epilepsi juga dapat memiliki dampak psikologis yang signifikan pada individu, seperti depresi, kecemasan, dan stigmatisme sosial. Beberapa penderita epilepsi mungkin mengalami keterbatasan dalam aktivitas sehari-hari dan mobilitas, terutama jika mereka memiliki kejang yang sering atau berat.

Menurut data dari WHO (*World Health Organization*) hingga tahun 2021, diperkirakan ada sekitar 50 juta orang di seluruh dunia yang menderita epilepsi. Dalam kasus epilepsi yang tidak terkontrol dengan baik, risiko kematian dapat meningkat, terutama akibat kejang yang berat atau komplikasi medis yang terkait dengan gangguan ini. Namun, perlu diingat bahwa dengan pengobatan yang tepat, sebagian besar individu dengan epilepsi dapat mengontrol kejang mereka dan memiliki kualitas hidup yang baik. Oleh karena itu, upaya global terus dilakukan untuk meningkatkan pemahaman tentang epilepsi, menyediakan perawatan yang sesuai, memberikan dukungan psikososial, dan mengurangi stigma sosial sekitar penyakit ini, sehingga individu yang terkena dapat hidup dengan lebih baik dan aman.

Epilepsi adalah sebuah masalah kesehatan global yang memengaruhi jutaan orang di seluruh dunia, dan pengelolaan penyakit ini memerlukan deteksi dini yang akurat dan pengobatan yang tepat. Hingga saat ini epilepsi dibagi berdasarkan jenis

kejangnya yaitu kejang parsial dan kejang umum. Untuk diagnosis penyakit epilepsi dapat dilakukan dengan hal manual seperti cara berjalan, otot ataupun dengan pengukuran sinyal otak seperti *Magneto-Encephalography* (MEG), *functional Magnetic Resonance Imaging* (fMRI), dan *electroencephalogram* (EEG)

Dengan beberapa metode di atas metode penggunaan EEG dirasa lebih tepat dikarenakan EEG dapat dilakukan dengan lebih murah serta memiliki resolusi temporal yang lebih tinggi dibandingkan dengan MEG dan fMRI. EEG sendiri adalah metode pengukuran aktifitas listrik spontan dari otak yang diperoleh dengan cara menembakkan sinyal listrik ke neuron dalam otak. EEG mempunyai resolusi temporal yang tinggi dan dapat mendeteksi segala perubahan pada aktifitas otak, hasil akhir dari EEG ialah sebuah gambar yang menunjukkan perubahan sinyal. Sinyal EEG berbentuk gelombang elektrik yang ukurannya sangat kecil.

Agar mempermudah membaca EEG dapat menggunakan metode *Neural Network* yang dikembangkan menjadi *Convolutional Neural Network* (CNN). Metode *Neural Network* ini sudah pernah dilakukan dalam penelitian sebelumnya, hanya saja pada penelitian tersebut menggunakan CHB-MIT dataset. Pada penelitian itu dihasilkan tingkat akurasi mencapai 97.35%. [8]

Berdasarkan latar belakang di atas, maka dilakukan Penelitian Proyek Akhir yang diberi judul “Klasifikasi sinyal normal dan kejang menggunakan *Convolutional Neural Network*”. Hasil dari penelitian ini diharapkan mendapatkan model tingkat akurasi yang lebih baik dengan menggunakan data yang berbeda. Klasifikasi epilepsi pada EEG dibagi menjadi 3 kelas yaitu normal, ictal dan interictal. Selain itu juga diharapkan mampu membantu dalam proses deteksi penyakit epilepsi dengan akurat dan mendetail.

## 1.2 Tujuan dan Manfaat

Adapun Tujuan dari proyek akhir ini yaitu:

1. Mengembangkan Model Klasifikasi 1D CNN yang Efektif: Tujuan utama dari penelitian ini adalah mengembangkan model jaringan saraf konvolusi dalam dimensi 1 (1D CNN) yang efektif dalam mengklasifikasikan sinyal EEG menjadi dua kategori utama, yaitu sinyal epilepsi dan sinyal normal.
2. Meningkatkan Diagnosis Dini Epilepsi: Penelitian ini bertujuan untuk meningkatkan kemampuan diagnosis dini epilepsi melalui otomatisasi

analisis sinyal EEG. Dengan model yang akurat, diagnosis dapat dilakukan lebih awal, memungkinkan intervensi dan perawatan yang lebih efektif.

3. Pengembangan Alat Bantu Medis: Proyek ini bertujuan untuk menghasilkan alat bantu medis yang dapat digunakan oleh profesional kesehatan untuk mendukung proses diagnosis epilepsi. Hal ini akan membantu dalam meningkatkan efisiensi diagnosis dan perawatan pasien.
4. Pengembangan Pemahaman Terhadap Data EEG: Tujuan lainnya adalah untuk mengembangkan pemahaman yang lebih baik tentang karakteristik data EEG dan cara terbaik untuk mengolahnya menggunakan teknologi CNN. Ini dapat berkontribusi pada penelitian medis lebih lanjut di masa depan.

Adapun manfaat dari proyek akhir ini yaitu:

1. Diagnosis Dini Epilepsi: Penelitian ini akan memberikan manfaat besar dalam meningkatkan diagnosis dini epilepsi. Pasien yang terkena epilepsi dapat mendapatkan perawatan yang lebih cepat, yang dapat mengurangi dampak serangan epilepsi dan meningkatkan kualitas hidup mereka.
2. Otomatisasi Analisis EEG: Model CNN yang dikembangkan dapat membantu mengotomatisasi proses analisis sinyal EEG. Hal ini akan menghemat waktu dan usaha profesional kesehatan dalam menganalisis data EEG yang kompleks.
3. Perbaikan Manajemen Pasien: Dengan diagnosis yang lebih cepat dan akurat, manajemen pasien epilepsi dapat ditingkatkan. Profesional medis akan dapat merencanakan perawatan yang lebih tepat sasaran dan memantau respons pasien dengan lebih baik.
4. Kontribusi pada Penelitian Medis: Penelitian ini dapat berkontribusi pada pemahaman lebih lanjut tentang penggunaan teknologi CNN dalam pemrosesan sinyal EEG dan dapat mendorong penelitian medis lebih lanjut dalam bidang ini.

### **1.3 Rumusan Masalah**

Adapun rumusan masalah dalam penelitian ini yaitu:

1. Bagaimana cara untuk membuat sistem deteksi epilepsi menggunakan CNN?
2. Bagaimana memunculkan karakter sinyal berdasarkan sinyal EEG 1 dimensi?
3. Bagaimana tingkat akurasi yang dihasilkan dari sistem yang dibuat?
4. Apakah pendekatan ini memiliki potensi untuk menjadi alat bantu dalam diagnosis

dini epilepsi, dan bagaimana dampaknya terhadap perawatan dan manajemen pasien dengan epilepsi?

#### 1.4 Batasan Masalah

Dalam Proyek Akhir ini, dilakukan pembatasan masalah sebagai berikut

1. Data sampel yang digunakan berdasarkan Bonn dataset
2. Data sinyal yang digunakan terbatas pada sinyal EEG
3. Pada penelitian ini menganalisis epilepsi dalam kondisi normal, ictal, dan interictal
4. Sistem ini dijalankan dengan google colab menggunakan bahasa python
5. Optimasi yang digunakan adalah *Adam optimizers*

#### 1.5 Metodologi

Metodologi pada penelitian ini, sebagai berikut:

##### 1. Studi Literatur

Studi literatur dilakukan dengan mengumpulkan literatur-literatur dan kajian-kajian yang berkaitan dengan permasalahan yang ada pada penelitian Proyek akhir ini, baik berupa buku referensi, artikel, maupun *e-journal* yang berhubungan dengan *Convolutional Neural Network*, Epilepsi, dan *Google Colab*.

##### 2. Perancangan sistem

Perancangan sistem dilakukan dengan cara menganalisa data studi literatur dan merancang sistem klasifikasi penyakit epilepsi menggunakan algoritma *Convolutional Neural Network* menggunakan bahasa pemrograman python

##### 3. Pengembangan dan Pelatihan Model CNN

Hal yang dilakukan adalah penentuan arsitektur CNN yang sesuai untuk tugas klasifikasi EEG. Ini mencakup jumlah lapisan konvolusi, ukuran kernel, dan lapisan tersembunyi lainnya.

##### 4. Pengujian sistem

Evaluasi model yang telah dilatih menggunakan data pengujian independen dengan metrik seperti akurasi, F1-score, sensitivitas, dan spesifisitas.

##### 5. Dokumentasi

Membuat dokumentasi atau laporan kesimpulan akhir dari analisa dan pengujian dalam bentuk Buku Proyek Akhir.

## 1.6 Sistematika Penulisan

Dalam penulisan Proyek Akhir terdiri atas lima bab, dengan keterangan sebagai berikut:

### **BAB I      PENDAHULUAN**

Pada bab ini berisi latar belakang, rumusan masalah, tujuan dan manfaat, batasan masalah, metodologi penelitian, serta sistematika penulisan.

### **BAB II     DASAR TEORI**

Pada bab ini membahas tentang teori pendukung pengerjaan proyek akhir, seperti konsep Neural Network, struktur Convolutional Neural Network, dan lain sebagainya.

### **BAB III    PERANCANGAN SISTEM**

Pada bab ini membahas tentang deskripsi sistem, proses pengerjaan sistem, dan hasil pengujian beberapa sistem.

### **BAB IV    ANALISIS HASIL PENGUJIAN**

Pada bab ini membahas hasil pengujian.

### **BAB V     PENUTUP**

Pada bab ini membahas tentang kesimpulan dari pengerjaan Proyek Akhir dan saran untuk pembaca yang akan mengambil penelitian dengan topik yang sama.

## **BAB II**

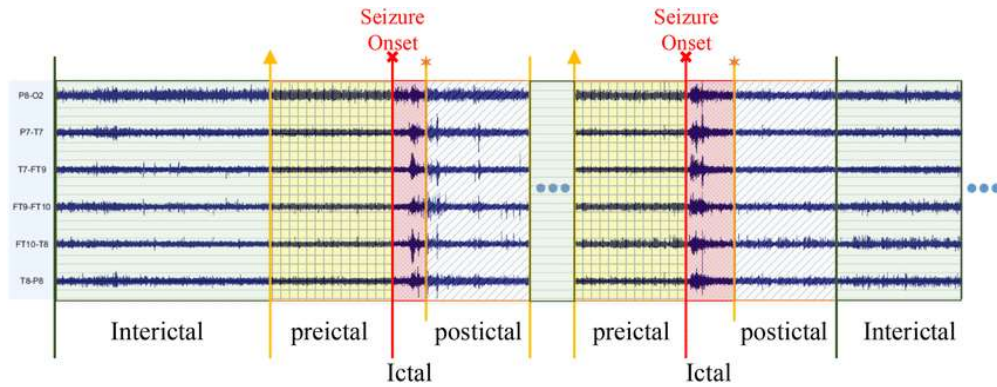
### **DASAR TEORI**

#### **2.1 Epilepsi**

Epilepsi adalah gangguan neurologis yang menandai dirinya dengan kejadian berulang yang tidak terkendali dari aktivitas listrik abnormal dalam otak, yang dapat menyebabkan serangan atau kejang. Gangguan ini adalah salah satu jenis gangguan neurologis yang paling umum dan dapat memengaruhi individu dari segala usia. Kejadian epilepsi sering kali disertai dengan serangan atau kejang yang dapat mempengaruhi berbagai fungsi tubuh. Kejang dapat terjadi dalam berbagai bentuk, mulai dari gerakan tubuh yang tidak terkendali hingga perubahan kesadaran dan perilaku.

Penyebab epilepsi dapat bervariasi, dengan sebagian besar kasus tidak memiliki penyebab yang jelas dan dikenal sebagai epilepsi idiopatik. Namun, beberapa kasus dapat disebabkan oleh faktor tertentu seperti cedera kepala, infeksi otak, tumor otak, kelainan perkembangan otak, atau masalah perinatal saat lahir. Gejala epilepsi melibatkan terutama kejang, yang bisa berupa kejang tonik-klonik (beku dan berkedut), kejang tonik (ketegangan otot), kejang klonik (gerakan berulang), kejang absen (kehilangan kesadaran yang singkat), dan lainnya. Pemeriksaan klinis, *Elektroensefalografi* (EEG), dan pencitraan otak seperti MRI digunakan dalam proses diagnostik.

Pengobatan epilepsi melibatkan penggunaan obat-obatan antiepilepsi (AEDs) untuk mengendalikan kejang. Dalam beberapa kasus yang resisten terhadap pengobatan, pembedahan otak atau terapi lainnya seperti stimulasi saraf vagus dapat dipertimbangkan. Dengan pengobatan yang tepat dan dukungan yang memadai, banyak individu dengan epilepsi dapat memiliki kualitas hidup yang baik dan produktif.



**Gambar 2.1** Sinyal EEG untuk fase berbeda pada epilepsi [9]

Sinyal EEG dalam fase yang berbeda, termasuk interictal, preictal, ictal, dan postictal. Fase ictal adalah periode terjadinya epilepsi. Fase sebelum dan sesudah fase ictal adalah fase preictal dan postictal. Terakhir, fase interiktal adalah periode di antara kejang. [9]

## 2.2 Neural Network

Jaringan saraf tiruan, atau neural network, adalah model matematika yang terinspirasi oleh struktur dan fungsi jaringan saraf biologis dalam otak manusia. Ini adalah komponen kunci dalam bidang kecerdasan buatan (artificial intelligence) yang digunakan untuk memproses data dan mengambil keputusan dengan cara yang meniru kemampuan belajar dan adaptasi manusia.

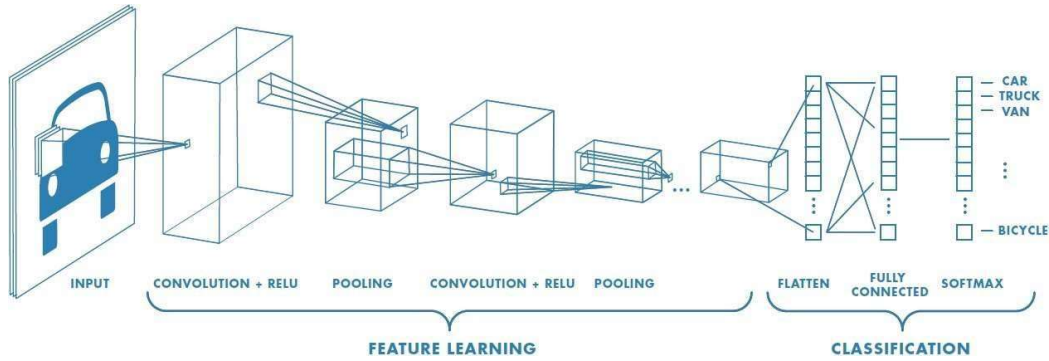
## 2.3 Convolutional Neural Network

*Convolutional Neural Network* (CNN), atau Jaringan Saraf Konvolusi, adalah jenis arsitektur jaringan saraf tiruan yang dirancang khusus untuk mengolah data yang memiliki struktur spasial, seperti citra dan data berdimensi tinggi seperti sinyal waktu-seri. CNN terinspirasi oleh cara manusia mengenali pola visual dalam gambar, dengan penggunaan lapisan konvolusi yang mampu mengidentifikasi fitur-fitur berbeda dalam data input. Arsitektur CNN terdiri dari beberapa lapisan yang meliputi lapisan konvolusi, lapisan aktivasi (seperti ReLU), dan lapisan pooling untuk mengekstraksi fitur-fitur penting dari data input. CNN juga umumnya memiliki lapisan fully connected di akhir untuk klasifikasi atau regresi. Keunggulan CNN adalah kemampuannya untuk belajar fitur-fitur hierarkis dari



data tanpa adanya ekstraksi fitur manual.

Oleh karena itu, CNN telah membawa kemajuan signifikan dalam berbagai aplikasi, termasuk pengenalan gambar, deteksi objek, pengolahan citra medis, serta pemrosesan sinyal, seperti yang mungkin ditemui dalam proyek klasifikasi sinyal epilepsi dan normal menggunakan CNN dalam 1 dimensi.



Sumber : (<https://saturncloud.io/blog/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/>)

**Gambar 2.2** Convolutional Neural Network (CNN)

### 2.3.1 Input

*Input* berfungsi untuk menampung nilai piksel dari citra yang menjadi masukan. Untuk citra dengan ukuran  $150 \times 150$  dengan 3 channel warna, RGB (*Red, Green, Blue*) maka yang akan menjadi masukan adalah piksel array yang berukuran  $150 \times 150 \times 3$ . Setiap warna pada tiap pixel gambar mempunyai nilai antara 0 sampai 255.

### 2.3.2 Convolution layer

*Convolution layer* adalah inti dari CNN. *Convolution layer* menghasilkan citra baru yang menunjukkan fitur dari citra *input*. Dalam proses tersebut, *Convolution layer* menggunakan filter pada setiap citra yang menjadi masukan. *Filter* pada layer ini berupa array 2 dimensi bisa berukuran  $5 \times 5$ ,  $3 \times 3$  atau  $1 \times 1$ . Proses *Convolution layer* dengan menggunakan *filter* pada layer ini akan menghasilkan *feature map* yang akan ditambahkan *activation function*.

### 2.3.3 Pooling layer

Pooling layer berfungsi untuk menerima input dari activation function kemudian mengurangi jumlah parameter nya. Pooling layer juga biasa disebut subsampling atau downsampling yang mengurangi dimensi dari feature map tanpa menghilangkan informasi penting di dalamnya. Pooling layer yang biasa digunakan adalah Max Pooling dan Average Pooling. Max Pooling untuk

menentukan nilai maksimum tiap pergeseran filter, sementara Average Pooling akan menentukan nilai rata-rata nya.

#### **2.3.4 Fully connected layer**

*Feature map yang dihasilkan dari proses convolutional layer dan pooling layer masih berbentuk multidimensional array, sehingga harus melakukan “flatten” atau reshape feature map menjadi sebuah vector agar bisa digunakan sebagai input dari fully connected layer. Fully connected layer merupakan lapisan dimana semua neuron aktivitas dari lapisan sebelumnya terhubung semua dengan neuron di lapisan selanjutnya seperti halnya jaringan syaraf tiruan biasa. Setiap aktivitas dari lapisan sebelumnya perlu diubah menjadi data satu dimensi sebelum dapat dihubungkan ke semua neuron di fully connected layer.*

### **2.4 Electroencephalography (EEG)**

*Elektroensefalogram (EEG)* adalah teknik medis yang digunakan untuk merekam dan memonitor aktivitas listrik di dalam otak manusia. Ini adalah alat penting dalam bidang neurologi untuk mendiagnosis berbagai gangguan otak, termasuk epilepsi, gangguan tidur, cedera kepala, dan gangguan neurologis lainnya. Proses EEG melibatkan meletakkan serangkaian elektroda kecil pada kulit kepala pasien, yang terhubung ke perangkat perekam. Elektroda ini merekam sinyal listrik yang dihasilkan oleh aktivitas sel-sel otak. Data ini kemudian ditransformasikan menjadi grafik yang disebut EEG yang merekam pola aktivitas otak dalam bentuk gelombang otak.

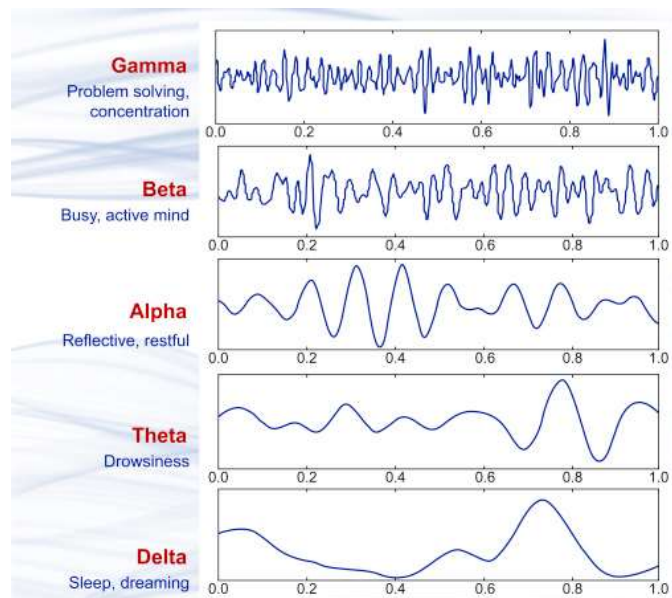
EEG dapat memberikan wawasan berharga tentang kondisi otak seseorang. Pola gelombang otak yang berbeda dapat mengindikasikan berbagai keadaan, termasuk tingkat kesadaran, aktivitas tidur, dan gangguan seperti serangan epilepsi. Ini juga digunakan dalam penelitian neurologis untuk memahami fungsi otak dalam berbagai konteks.

Dalam proyek akhir ini, EEG digunakan untuk merekam sinyal otak, yang kemudian dianalisis menggunakan *Convolutional Neural Network (CNN)* untuk mengklasifikasikan sinyal epilepsi dan normal. EEG adalah alat penting dalam pengukuran aktivitas otak dan memberikan wawasan yang sangat berharga dalam bidang medis dan penelitian neurologis.

Terdapat beberapa klasifikasi gelombang utama pada otak yang digunakan

sebagai diagnosis pengobatan dapat diperoleh dalam rentang frekuensi sebagai berikut:

1. Gelombang *Delta*, yaitu gelombang otak dengan frekuensi  $< 4$  Hz. Gelombang *delta* muncul saat dalam aktivitas yang rendah dan tidur lelap. Biasa terjadi pada orang dewasa saat tidur tenang dan pada bayi.
2. Gelombang *Tetha*, yaitu gelombang otak dengan frekuensi  $4 - 7$  Hz. Gelombang *tetha* muncul saat dalam kondisi yang merasakan kantuk dan saat aktivitas otak yang rendah. Biasa terjadi pada anak kecil beranjak remaja.
3. Gelombang *Alpha*, yaitu gelombang otak dengan frekuensi  $7 - 13$  Hz. Gelombang *alpha* muncul saat keadaan rileks dan saat memejamkan mata. Gelombang ini juga dikaitkan dengan kontrol penghambatan waktu pada aktivitas otak.
4. Gelombang *Beta*, yaitu gelombang otak dengan frekuensi  $13 - 35$  Hz. Gelombang *beta* muncul saat dalam keadaan sadar, aktivitas otak dan perilaku motorik. Gelombang *beta* terekam saat keadaan mata terbuka. Gelombang *beta* juga terjadi saat berpikir aktif, fokus, tingkat kewaspadaan yang tinggi serta saat keadaan cemas.
5. Gelombang *Gamma*, yaitu gelombang otak dengan frekuensi  $30 - 100$  Hz. Gelombang *gamma* muncul saat dalam tekanan mental yang tinggi juga saat keadaan sangat panik saat kondisi sadar.



**Gambar 2.3** Sinyal EEG

## 2.5 Google Colab



Sumber : (<https://www.detik.com/bali/berita/d-6476973/mengenal-google-colab-cara-menggunakan-dan-keuntungannya>)

**Gambar 2.4** Logo Google Colab

Google Colab adalah sebuah platform komputasi berbasis cloud yang disediakan oleh Google. Ini merupakan layanan yang memungkinkan pengguna untuk membuat, menjalankan, dan berbagi kode Python serta proyek data *science* secara online tanpa memerlukan perangkat keras khusus atau konfigurasi yang rumit. Berikut adalah beberapa poin penting tentang *Google Colab*:

1. Akses Gratis: *Google Colab* tersedia secara gratis bagi siapa saja yang memiliki akun Google. Dapat juga diakses melalui browser web tanpa perlu menginstal perangkat lunak tambahan di komputer.
2. Lingkungan Jupyter Notebook: *Google Colab* menggunakan lingkungan Jupyter Notebook yang sangat populer dalam pengembangan data science. Ini memungkinkan siapapun untuk menulis kode Python dalam sel-sel yang dapat dieksekusi secara interaktif, juga dapat menyisipkan teks, gambar, grafik, dan catatan dalam notebook.
3. Pemrosesan di Cloud: Yang membuat *Google Colab* sangat berguna adalah kemampuannya untuk menjalankan kode di cloud Google, yang berarti sumber daya komputasi yang kuat dapat diakses tanpa perlu memiliki perangkat keras berkekuatan tinggi. Ini sangat bermanfaat untuk pekerjaan yang memerlukan pengolahan data yang berat, seperti deep learning.
4. GPU dan TPU: *Google Colab* menyediakan akses ke GPU (Graphics Processing Unit) dan TPU (Tensor Processing Unit) untuk akselerasi komputasi. Ini sangat berguna untuk melatih model deep learning yang memerlukan daya komputasi tinggi.
5. Pustaka Siap Pakai: Colab telah dilengkapi dengan banyak pustaka Python yang umum digunakan untuk data science dan machine learning, seperti TensorFlow,

PyTorch, NumPy, dan sebagainya. Proses untuk menggunakan dan *impor* pustaka akan menjadi lebih mudah.

6. Kolaborasi dan Berbagi: Dapat berbagi notebook Colab dengan orang lain dan berkolaborasi secara online. Ini memudahkan tim dalam mengerjakan proyek data science bersama-sama.
7. Penyimpanan Google Drive: Dapat menyimpan notebook Colab secara langsung di Google Drive, yang memudahkan penyimpanan dan pengelolaan proyek Anda.
8. Pilihan untuk Menjalankan di CPU: Dapat juga memilih untuk menjalankan kode di CPU jika tidak memerlukan GPU atau TPU, yang berguna jika ingin menghemat sumber daya komputasi.

## 2.6 Bonn Dataset

Bonn EEG Dataset University adalah sumber daya penting dalam penelitian neurologi, khususnya dalam analisis sinyal otak menggunakan *Electroencephalogram* (EEG). Dataset ini dikumpulkan di Universitas Bonn, Jerman, dan memiliki karakteristik khusus yang membuatnya bernilai tinggi dalam pengembangan algoritma untuk analisis EEG serta deteksi dan klasifikasi kejang epilepsi.

## 2.7 Cross Validation

5-fold	Dataset				
Experiment 1	Test	Train	Train	Train	Train
Experiment 2	Train	Test	Train	Train	Train
Experiment 3	Train	Train	Test	Train	Train
Experiment 4	Train	Train	Train	Test	Train
Experiment 5	Train	Train	Train	Train	Test

Sumber : (<https://www.oreilly.com/library/view/mastering-predictive-analytics/9781789617740/8601ca16-be9f-4e93-81ba-12013e44c6b0.xhtml>)

**Gambar 2.5** K-Fold Cross Validation

K-Fold Cross-Validation adalah sebuah teknik yang digunakan dalam machine learning untuk mengevaluasi performa model dengan cara yang obyektif dan informatif. Metode ini sangat penting dalam menghindari overfitting, di mana model terlalu cocok dengan data pelatihan dan kurang mampu dalam menggeneralisasi pada data yang belum pernah dilihat sebelumnya.

Proses K-Fold Cross-Validation dimulai dengan membagi dataset yang

tersedia menjadi  $K$  subset yang sebanding dalam ukuran, dengan  $K$  adalah bilangan bulat yang dipilih (misalnya,  $K = 5$ ). Selama setiap iterasi, salah satu dari  $K$  fold digunakan sebagai data pengujian, sedangkan  $K-1$  fold lainnya digunakan sebagai data pelatihan. Model machine learning akan dilatih pada data pelatihan dan diuji pada data pengujian.

Selanjutnya, metrik evaluasi seperti akurasi, F1-score, atau yang lainnya dihitung untuk setiap iterasi, memberikan gambaran tentang seberapa baik model berkinerja pada setiap fold pengujian. Hasil metrik evaluasi dari masing-masing iterasi ini kemudian diintegrasikan menjadi satu nilai rata-rata atau statistik lainnya. Ini memberikan perkiraan yang lebih stabil tentang performa model daripada hanya mengandalkan hasil dari satu percobaan pembagian data.

Manfaat utama dari K-Fold Cross-Validation adalah membantu memastikan bahwa model yang dikembangkan mampu menggeneralisasi dengan baik pada data yang belum pernah dilihat sebelumnya. Selain itu, teknik ini juga membantu dalam pemilihan model yang paling cocok untuk tugas tertentu berdasarkan metrik evaluasi yang obyektif. Dengan menggunakan K-Fold Cross-Validation, kita dapat membuat evaluasi performa model yang lebih kredibel dan dapat diandalkan dalam pengembangan solusi machine learning.

## **2.8 Bahasa Pemrograman Python**

Python merupakan bahasa pemrograman tingkat tinggi (high-level language) yang berarti syntax python mendekati bahasa manusia sehingga dapat dengan mudah dipahami. Python telah dikembangkan sejak tahun 1980 oleh Guido van Rossum dan masih dikembangkan hingga saat ini

Python adalah salah satu bahasa pemrograman yang paling populer di dunia, digunakan dalam berbagai domain seperti pengembangan web, pengolahan data, kecerdasan buatan, analisis data, ilmu pengetahuan data, dan banyak lagi. Kelebihan Python meliputi komunitas yang besar dan aktif, banyaknya pustaka dan alat pengembangan yang tersedia, serta kemudahan dalam mengembangkan prototipe dan proyek-proyek kompleks.

Python memiliki beberapa karakteristik yang membuatnya menonjol. Pertama, Python memiliki sintaksis yang mudah dibaca dan dekat dengan bahasa manusia, yang membuatnya sangat sesuai untuk pemula dan pemrogram yang berpengalaman.