

# BAB I

## PENDAHULUAN

### A. Latar Belakang

Sebelum berkembangnya teknologi pada bidang kesehatan, terdapat banyak kendala yang dihadapi petugas kesehatan seperti dalam hal pemantauan tingkat stres dan kondisi kesehatan pasiennya masih menggunakan cara tradisional. Sebagai contoh ketika petugas kesehatan akan melakukan pemantauan tingkat stres dari pasiennya maka pasien akan diberikan sebuah pertanyaan psikologis yang kemudian petugas harus memberikan bobot secara manual dan kemudian baru dapat menentukan hasilnya. Jika terus menggunakan cara tradisional ini, maka dapat memperlambat kinerja dari petugas kesehatan. Lalu, jika ada seorang pasien yang tinggalnya sangat jauh dari pusat kesehatan, maka ini akan menambah sulit bagi petugas untuk memantau kesehatan dari pasiennya.

Seiring dengan berkembangnya teknologi pada bidang kesehatan, kini petugas kesehatan dapat memantau kesehatan pasiennya dari jarak jauh. Petugas kesehatan dapat memanfaatkan teknologi *Wearable Sensor* seperti *Empatica E4 (Wristband Sensor)*, *Polar H10 (Belt/Chest strap Sensor)*, *Garmin HRM-Pro (Chest strap Sensor)*, dan *RespiBAN (Chest strap Sensor)*. Dengan menggunakan berbagai *Wearable Sensor* seperti yang telah disebutkan diatas, petugas kesehatan dapat memantau kesehatan pasien yang bermukim jauh dari pusat kesehatan. Pasien dapat

menggunakan *Wearable Sensor* ketika petugas kesehatan meminta untuk melakukan pengecekan kesehatan saja, tidak perlu untuk dipakai setiap saat.

Data kesehatan pasien yang didapat menggunakan *Wearable Sensor* kemudian harus dianalisa kembali. Untuk dapat menganalisa data tersebut terdapat beberapa permasalahan: pertama, ukuran data yang besar sehingga butuh sumber komputasi yang besar dan juga koneksi internet yang stabil untuk mengupload data tersebut. Kedua, karena data kesehatan itu salah satu data yang bersifat pribadi, sehingga riskan apabila di-upload ke server secara mentah-mentah (“Personal Data | General Data Protection Regulation (GDPR),” n.d.). Penelitian yang dilakukan oleh Kajornkasirat, Chanapai, & Hnusuwan (2018), penelitian ini merupakan salah satu contoh penelitian yang masih menggunakan cara ‘tradisional’ dalam memproses data mentah yang diambil dari *Wearable Sensor* yaitu dengan langsung mengupload data mentah tersebut ke *cloud server* sehingga riskan jika *cloud server* mengalami kebocoran data.

*Edge Computing* merupakan salah satu alternatif yang dapat digunakan untuk melakukan proses pengolahan data yang telah diambil dari *Wearable Sensor* dengan cara menerapkan sebuah perangkat yang sangat dekat dengan sumber data dihasilkan. Dengan memanfaatkan *edge devices* untuk melakukan proses komputasi secara lokal tanpa harus ada proses pengiriman data mentah dari sensor yang kemudian di-upload ke server. Dengan memanfaatkan metode *Federated Learning* untuk melakukan pemrosesan data mentah pada *edge devices* yang memiliki

keunggulan tidak memerlukan koneksi internet yang cepat dan juga membutuhkan *latency* yang rendah, serta efisien dari segi sumber daya komputasi.

Salah satu penelitian yang menerapkan metode *Federated Learning* adalah oleh Hard dkk (2018) yaitu pada produk *Google Keyboard* untuk memprediksi kata yang sering digunakan oleh penggunanya. Hal ini bertujuan agar tidak adanya data pribadi yang diketikkan oleh penggunanya ter-upload ke server pusat Google. Dengan demikian proses latihan model pemrosesan Bahasa alami yang dibuat oleh Google dapat dilakukan pada perangkat setiap penggunanya. Namun, pihak Google belum merilis bagaimana cara untuk melakukan *deploying* model *Federated Learning* langsung pada *edge devices*.

Ada beberapa alasan mengapa penting untuk melakukan pemrosesan data yang bersifat sensitif pada *edge devices* di antaranya adalah penghematan biaya, keamanan data, dan performa. Penghematan biaya karena jika terdapat banyak *wearable sensor* yang digunakan, kemudian setiap sensor mengirimkan data ke *cloud server* secara *real-time* maka hal ini akan menambah biaya pengiriman data melalui jaringan seluler. Terlebih karena data yang dikirim berukuran besar, tentu membutuhkan kecepatan internet yang stabil dan rendah *latency*. Untuk itu, dengan tetap melakukan pemrosesan data pada *edge devices* termasuk menghemat biaya. Keamanan data, dengan tetap melakukan pemrosesan pada secara lokal, maka akan semakin kecil resiko dari kebocoran data pribadi pasien yang telah diambil dari *wearable sensor*. Jika terjadi suatu anomali dapat dengan mudah terdeteksi karena

data yang bersifat lokal, tidak perlu mengecek pada *cloud server*. Masalah performa adalah sesuatu yang harus diperhatikan dalam penelitian semacam ini, karena dengan besar dan banyaknya data yang harus dilatih maka dibutuhkan performa yang bagus mulai dari koneksi internet hingga kemampuan dari sumber daya yang harus mumpuni. Sehingga dengan melatih data pada *edge devices* dapat meningkatkan performa dari proses pelatihan data, karena data bersifat lokal dan tidak terpengaruh akan *latency* serta kemampuan *edge devices* yang cukup mumpuni untuk melakukan pemrosesan data secara *multi-threading* (“Training at the Network Edge with AI: 3 Key Benefits - RTInsights,” n.d.).

Berdasarkan beberapa permasalahan yang telah dibahas mengenai pengolahan data terutama data kesehatan, maka penelitian tentang pemantauan kesehatan menggunakan *IoT* dan metode *Federated Learning* ini dapat menjadi alternatif dalam menyelesaikan masalah pengolahan data kesehatan. Penelitian ini menggunakan data yang diambil menggunakan sensor *RespiBAN (Chest-strap Sensor)* dan *Empatica E4 (Wristband Sensor)*. Namun, hanya data dari sensor *Empatica E4 (Wristband Sensor)* yang akan digunakan pada penelitian ini. Sensor *Empatica E4* memiliki bentuk yang menyerupai *smartwatch* yang diletakkan juga pada pergelangan tangan, sehingga memudahkan pasien yang bermukim jauh dari pusat kesehatan ketika melakukan pengambilan data kesehatan. Berbeda dengan sensor *RespiBAN* yang membutuhkan banyak kabel yang dipasang pada dada pasien, sehingga menyulitkan pasien yang bermukim jauh dari pusat kesehatan ketika akan dilakukan pengecekan kondisi kesehatan (Indikawati & Winiarti, 2020).

## B. Identifikasi Masalah

Berdasarkan uraian dari latar belakang, maka dapat ditentukan identifikasi masalahnya sebagai berikut:

1. Data yang telah diambil dari *wearable sensor* merupakan data yang sensitif dan bersifat privasi.
2. Masih terdapat penelitian yang melatih data yang bersifat sensitif langsung ke *cloud server*.
3. Masih maraknya penelitian yang melakukan pemodelan data tanpa memperdulikan legalitas perihal keamanan data.

## C. Batasan Masalah

Agar penelitian ini tidak meluas, perlu adanya Batasan masalah. Adapun batasan masalahnya adalah sebagai berikut:

1. Penelitian ini akan mensimulasikan metode *Federated Learning* dalam proses pemodelan data dari seluruh subjek terhadap kelas klasifikasi (*baseline, stress, amusement*).
2. Penelitian ini juga menggunakan model klasifikasi *Random Forest* sebagai *global model* pada penerapan simulasi *Federated Learning*.
3. *Dataset* yang digunakan dalam penelitian ini adalah *Wearable Stress and Affect Detection* yang diambil menggunakan sensor *RespiBAN* dan

*Empatica E4* yang merupakan sensor untuk bagian dada dan pergelangan tangan.

#### **D. Rumusan Masalah**

Berdasarkan identifikasi dan batasan masalah diatas, maka masalah yang dapat dirumuskan adalah sebagai berikut:

1. Bagaimana melakukan implementasi klasifikasi *Random Forest* dalam dataset WESAD untuk menentukan kondisi dari setiap subjek (*Baseline, Stress, Amusement*).
2. Bagaimana mensimulasikan *Federated Learning* dalam jumlah subjek sebanyak 7 pada WESAD?
3. Bagaimana menggunakan menggunakan *Leave one subject out* pada proses persiapan data?

#### **E. Tujuan Penelitian**

Adapun tujuan dari penelitian ini adalah:

1. Menerapkan metode klasifikasi *Random Forest* dalam memprediksi keadaan dari setiap partisipan (*Baseline, Stress, Amusement*).
2. Mencoba mensimulasikan metode *Federated Learning*.

## F. Manfaat Penelitian

Manfaat yang didapat pada penelitian ini antara lain adalah:

1. Penelitian ini akan menghasilkan model data untuk mendeteksi apakah seorang partisipan dalam keadaan *baseline, stress, amusement* berdasarkan data aktual kondisi psikologi partisipan yang terdapat pada *dataset* WESAD.
2. Mensimulasikan bagaimana *Federated Learning* bekerja pada data WESAD.