PipeLine Definida

Para realizarmos as classificações montamos a seguinte PipeLine:

1. Filtragem dos casos visualizados na Análise de Dados Faltantes.
2. Segregação em Treino (12 pessoas) e Test (5 pessoas) [Com K-Folds].
3. UnderSampling utilizando RandomUnderSampler().
4. Utilização de Interpolação para completar os dados faltantes nas series.
5. Utilização de Catch22 para extração de features das series temporais.

Com a PipeLine definida vamos visualizar os resultados de KNN k=5 e distância = dynamic time warping.

Utilizando K-Folds foi encontrado a seguinte segmentação entre treino e teste:

Interface gráfica do usuário, Aplicativo

Descrição gerada automaticamente

Resultados:

Texto

Descrição gerada automaticamente

Tabela

Descrição gerada automaticamente

Gráfico, Gráfico de mapa de árvore

Descrição gerada automaticamente

Visualizando os resultados, podemos perceber que não foi identificado nenhuma diferença entre um movimento Normal e um movimento Anormal. Podemos visualizar melhor quando olhamos para os resultados de uma das pessoas na base de teste:



Texto

Descrição gerada automaticamente com confiança média

Tabela

Descrição gerada automaticamente com confiança médiaGráfico, Gráfico de mapa de árvore

Descrição gerada automaticamente

Vamos começar visualizando alguns casos Negativos Verdadeiros (true = 0, predict = 0)

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Gráfico, Gráfico de linhas

Descrição gerada automaticamenteGráfico, Gráfico de linhas

Descrição gerada automaticamente

Gráfico, Gráfico de linhas

Descrição gerada automaticamenteGráfico, Gráfico de linhas

Descrição gerada automaticamenteInterface gráfica do usuário, Gráfico, Gráfico de linhas

Descrição gerada automaticamenteGráfico, Gráfico de linhas

Descrição gerada automaticamente

Como podemos visualizar grande parte dos casos são movimentos simples de se levantar e se sentar sem muitas alterações neste movimento, alguns desses movimentos apresentam essas mudanças bruscas de várias subidas e descidas. Da qual podem ser indicadores não mapeados de Fadiga, ou falha na coleta dos dados.

Agora vamos visualizar alguns casos de Verdadeiro Positivos (true = 1, predict = 1)

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Foto de pessoas

Descrição gerada automaticamente com confiança baixa

Gráfico

Descrição gerada automaticamente

Gráfico, Gráfico de linhas

Descrição gerada automaticamenteGráfico, Gráfico de linhas

Descrição gerada automaticamente

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Podemos visualizar que temos casos simples de indicadores de Fadiga, como Platô (muito tempo parado antes da subida ou descida) ou duplo (descida ou subida dupla) dos quais são visualmente diferentes dos casos normais. Mas também conseguimos visualizar alguns exemplos bem parecidos com os normais, dos quais também possuem as mesmas variações bruscas, podendo identificar outros casos de falha da coleta, ou anotações inconsistentes, que não identificam o mesmo problema como anormal.

Agora vamos visualizar alguns casos de Falso Positivos (True = 0, predict = 1)

Tela de computador com fundo branco

Descrição gerada automaticamente com confiança baixa

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Como podemos ver a maioria desses casos foram identificados pelo algoritmo como Anormal por causa das alterações rápidas ocorrida no começo ou fim do movimento. Esse comportamento pode ser explicado olhado para os casos apontados como Ritmo para uma das bases utilizadas no treino (Carlos):

Interface gráfica do usuário, Gráfico

Descrição gerada automaticamente com confiança média

Interface gráfica do usuário, Gráfico de linhas

Descrição gerada automaticamenteGráfico, Gráfico de linhas

Descrição gerada automaticamente

Todos esses casos acima foram identificados como Anormal (ritmo), mas isso não é um padrão utilizado por todas as bases, já que a Teresinha possuía esses mesmos casos, mas não era identificado como Anormal

Por fim, vamos visualizar alguns casos de falsos negativos (true = 1, predict = 0):

Gráfico, Gráfico de linhas

Descrição gerada automaticamenteGráfico, Gráfico de linhas

Descrição gerada automaticamente

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Os primeiros dois casos são casos de Platô, dos quais, são realmente platô considerando a movimentação da pessoa analisada, mas que não é um padrão para as outras pessoas utilizadas no treino. Já os dois últimos seguem no mesmo raciocínio dos casos anteriores, casos identificados como Anormal (ritmo) por causa das variações ocorridas no começo/fim do movimento, mas da qual não foi feita para toda a base

Podemos visualizar essas diferenças em formas de anotações antes mesmo do KNNGráfico, Gráfico de barras

Descrição gerada automaticamente

Aqui podemos ver que o Arnaldo e o Carlos possuem um número muito maior do que o observado nas outras pessoas de Anormalidades, então, enquanto parte da base recebeu o tratamento mais restrito de Anormalidade, outras já receberam tratamento mais irrestrito, identificando quase todo tipo de mudança como anormalidade.