1. **Descrição**

Este arquivo contém 13.910 medições de 16 sensores químicos expostos a 6 gases diferentes em vários níveis de concentração. Estas medições são utilizadas em simulações para compensação de deriva em uma tarefa de discriminação de seis gases em vários níveis de concentração.

O conjunto de dados foi coletado durante o período de janeiro de 2008 a fevereiro de 2011 (36 meses). Esses dados são divididos em 10 arquivos em lote com extensão '.dat'. O conjunto de dados foi coletado durante o período de janeiro de 2008 a fevereiro de 2011 (36 meses) em uma plataforma de distribuição de gás situada no Laboratório ChemoSignals no Instituto BioCircuits, Universidade da Califórnia em San Diego.

Detalhes da organização do conjunto de dados. Cada linha corresponde a meses que foram combinados para formar um lote:

IDs de mês de ID de lote

Lote 1 Meses 1 e 2

Lote 2 Meses 3, 4, 8, 9 e 10

Lote 3 Meses 11, 12 e 13

Lote 4 Meses 14 e 15

Lote 5 Mês 16

Lote 6 Meses 17, 18, 19 e 20

Lote 7 Mês 21

Lote 8 Meses 22 e 23

Lote 9 Meses 24 e 30

Lote 10 Mês 36

1. **Informação do Dataset**
   1. Resumo dos Dados

| **Number of Instances** | **Number of Attributes** | **Number of Missing Values(N/A)** | **Number of Classes** |
| --- | --- | --- | --- |
| 13910 | 129 | 0 | 6 |

A primeira variável nominal é uma variável de classe (alvo). As 128 variáveis restantes são todas do tipo numérico. Não há distinção entre dados de treinamento e dados de teste no arquivo de dados.

| **Name of gas** | **Ammonia** | **Acetaldehyde** | **Acetone** | **Ethylene** | **Ethanol** | **Toluene** |
| --- | --- | --- | --- | --- | --- | --- |
| ppmv | (50,1000) | (5,500) | (12,1000) | (10,300) | (10,600) | (10,100) |

Nome do gás das 6 classes e intervalo de níveis de concentração de cada gás

* 1. Quantidade de cada classe

| **1** | **2** | **3** | **4** | **5** | **6** |
| --- | --- | --- | --- | --- | --- |
| 2565 | 2926 | 1641 | 1936 | 3009 | 1833 |

1. **Atributos**

Por que um vetor de recursos de 128 elementos por medição? Os sensores de gás MOX normalmente descrevem uma mudança monotonicamente suave na condutância da camada de detecção devido aos processos de reação de adsorção/dessorção da substância de analito químico exposta. Representamos cada série temporal com um agregado de oito características que refletem a resposta do sensor. Em particular, consideramos dois tipos distintos de características na criação deste conjunto de dados: duas características de estado estacionário e seis características que refletem a dinâmica do sensor. As características de estado estacionário incluem a amplitude da mudança de resistência e seu valor normalizado. As características transitórias foram extraídas com base na média móvel exponencial (EMA) para refletir a dinâmica do sensor da porção transitória crescente/decadente das respostas do sensor. A transformada EMA avalia as porções crescentes/decrescentes da resistência do sensor considerando os valores máximo/mínimo de y[k] do seguinte filtro digital de primeira ordem: y[k] = (1-a)y[k-1] + uma(x[k] - x[k-1])

onde 0 < α < 1 é o parâmetro de suavização do filtro e x[k] é o valor adquirido no tempo k. Como diferentes valores de α fornecem diferentes valores de características e diferentes informações da resposta transitória, calculamos o filtro EMA para três valores de α = 0,1, 0,01, 0,001 para os estágios ascendente e descendente. Portanto, cada um dos 16 sensores utilizados no estudo contribui com 8 características, produzindo assim um vetor de características de 128 elementos por medição.