**Caracterização Produção PV**

**Lista de features utilizadas:**

|  |  |  |
| --- | --- | --- |
| Nome da Feature | Interpretação | Unidades |
| Lag51 | Produção PV 51h antes da hora a prever | kW |
| Lag50 | Produção PV 50h antes da hora a prever |
| Lag49 | Produção PV 49h antes da hora a prever |
| Lag48 | Produção PV 48h antes da hora a prever |
| Lag27 | Produção PV 27h antes da hora a prever |
| Lag26 | Produção PV 26h antes da hora a prever |
| Lag25 | Produção PV 25h antes da hora a prever |
| Lag24 | Produção PV 24h antes da hora a prever |
| Lag3 | Produção PV 3h antes da hora a prever |
| Lag2 | Produção PV 2h antes da hora a prever |
| Lag1 | Produção PV 1h antes da hora a prever |
| Temperatura\_24h | Temperatura registada 24h antes da hora a prever | ºC |
| Temperatura\_Prev | Temperatura prevista para a hora a prever |

Considerei o período de 17/1/2022 até 31/05/2023 como treino do modelo (cerca de 70% dos dados) e de 01/06/2023 até 31/12/2023 para teste.

Método de normalização utilizado: min-max

**Método de otimização**

*Simulated annealing* é uma das meta-heurísticas mais antigas e uma das suas principais características é a sua capacidade de conseguir escapar a mínimos ou máximos locais (consoante a função objetivo seja respetivamente de minimização ou de maximização). A ideia subjacente a esta característica do simulated annealing é a existência da probabilidade de aceitar soluções que eventualmente possam deteriorar a função objetivo de modo a afastar-se de possíveis mínimos (ou máximos) locais. A solução atual é, assim, substituída por uma nova solução gerada de forma aleatória na vizinhança, sendo que a extensão da busca é determinada por uma probabilidade de distribuição proporcional à “temperatura”. À medida que o algoritmo vai avançando e novas soluções vão sendo testadas, a temperatura vai diminuindo, ou seja, o espaço de soluções possíveis vai sendo cada vez menor. No entanto, de forma a evitar que o algoritmo se fixe num mínimo (máximo) local, existe sempre uma probabilidade de aumentar a temperatura (espaço de soluções) de forma a procurar novas soluções em outras regiões, procurando garantir que a função objetivo evite um mínimo (máximo) local e que se aproxime mais de um mínimo (máximo) absoluto.

Uma imagem com texto, captura de ecrã, Tipo de letra, design gráfico

Os conteúdos gerados por IA poderão estar incorretos.

Durante o processo do SA, em cada hora foram efetuados 30 ciclos (iterações). Em cada ciclo, 15 combinações aleatórias de parâmetros SVM foram testadas sendo que, no final de cada ciclo, a “temperatura” baixava de maneira a estreitar o intervalo de opções disponíveis para cada parâmetro. O parâmetro C (custo) foi testado entre 0 e 700, o Epsilon foi testado entre 0.00001 e 3 e o gamma foi testado entre 0.0001 e 10. A função objetivo definida para este processo foi baseada no RMSE gerado para o subconjunto de teste.

**Otimização dos parâmetros C, gamma e epsilon para as 15h**

Uma imagem com texto, diagrama, file, Gráfico

Os conteúdos gerados por IA poderão estar incorretos.

Nesta figura, são notórias as variações dos parâmetros a ser testados e as oscilações na função objetivo, comprovando a probabilidade de aceitar piores soluções durante o processo, mas tendendo para a minimização do erro.

**Uma imagem com texto, captura de ecrã, quadrado, padrão

Os conteúdos gerados por IA poderão estar incorretos.Matriz de correlação entre as variáveis utilizadas e os Outputs**

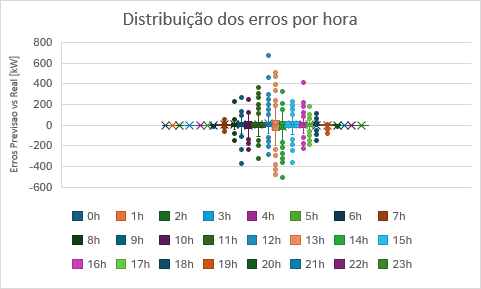
**Gráfico de Dispersão entre valores previstos e valores reais**

Uma imagem com texto, captura de ecrã, file, Gráfico

Os conteúdos gerados por IA poderão estar incorretos.

**Métricas de Erro**

|  |  |  |
| --- | --- | --- |
| **Metric** | **Train** | **Test** |
| MAE | 21,2 | 21,33 |
| MSE | 3129,79 | 2753,07 |
| RMSE | 55,94 | 52,47 |
| SSE | 1,78E+13 | 3,62E+12 |
| CV\_RMSE (%) | 36,28 | 32,49 |
| WAPE (%) | 13,75 | 13,21 |



\*Para o período de teste