



UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
PROJETO DE INICIAÇÃO CIENTÍFICA E TECNOLÓGICA
UNIDADE ACADÊMICA DE ENGENHARIA ELÉTRICA

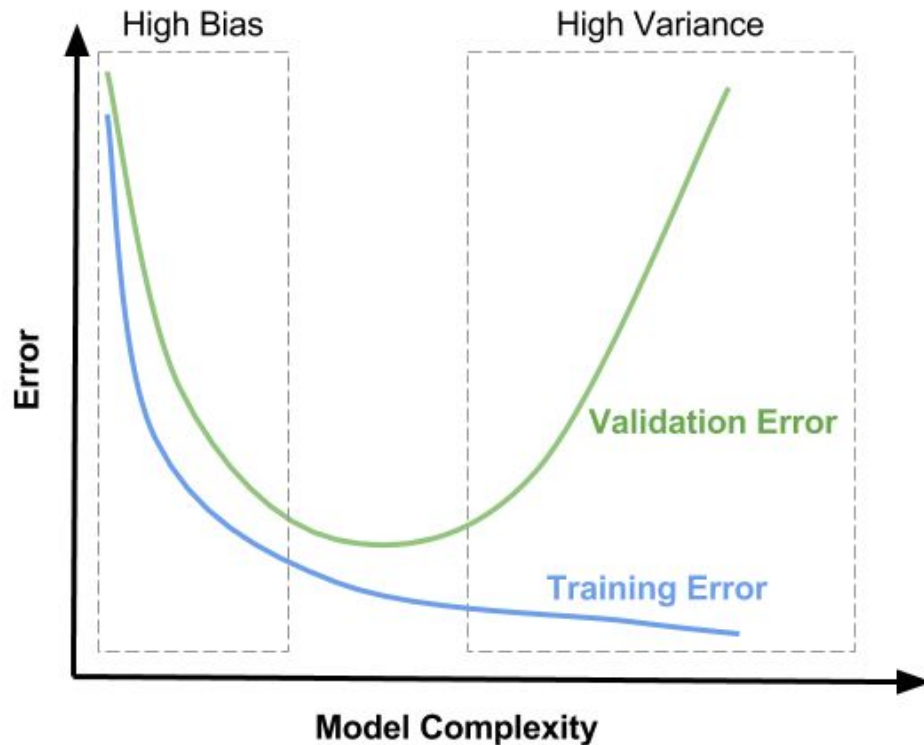
Viés vs Variância

Reunião 27/11/2020

Analizando Viés vs Variância

- Um dos principais problemas ao se analisar um modelo de aprendizagem de máquina é a distinção entre viés (bias) ou variância na contribuição de previsões ruins.
- Quando há um sobreajuste aos dados muito baixo (**underfitting**), estamos falando de um problema de alto viés (**high bias**).
- Em contrapartida, um sobreajuste excessivo aos dados (**overfitting**) é um problema de alta variância (**high variance**).

Analizando Viés vs Variância



Analizando Viés vs Variância

- Lembrando que o erro associado aos dados de treinamento e aos dados de validação cruzada são determinados por:

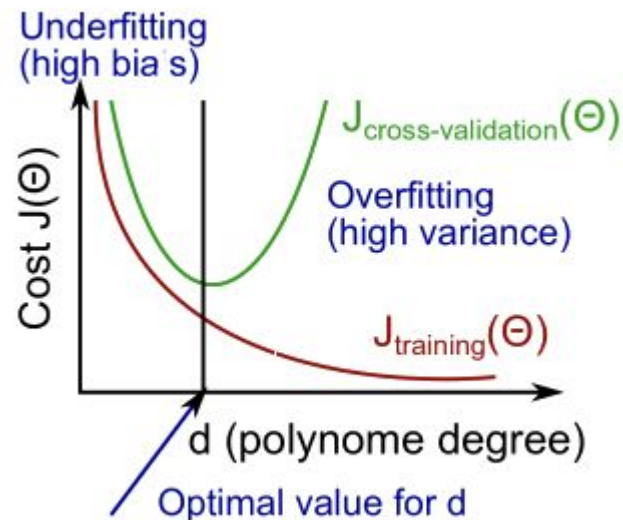
Training error: $\underline{J_{train}(\theta)} = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$

Cross validation error: $\underline{J_{cv}(\theta)} = \frac{1}{2m_{cv}} \sum_{i=1}^{m_{cv}} (h_{\theta}(x_{cv}^{(i)}) - y_{cv}^{(i)})^2$

Analizando Viés vs Variância

Viés Alto (Underfitting) : ambos $J_{train}(\Theta)$ e $J_{CV}(\Theta)$ será alto. Além disso, $J_{CV}(\Theta) \approx J_{train}(\Theta)$.

Alta variação (overfitting) : $J_{train}(\Theta)$ será baixo e $J_{CV}(\Theta)$ será muito maior do que $J_{train}(\Theta)$.



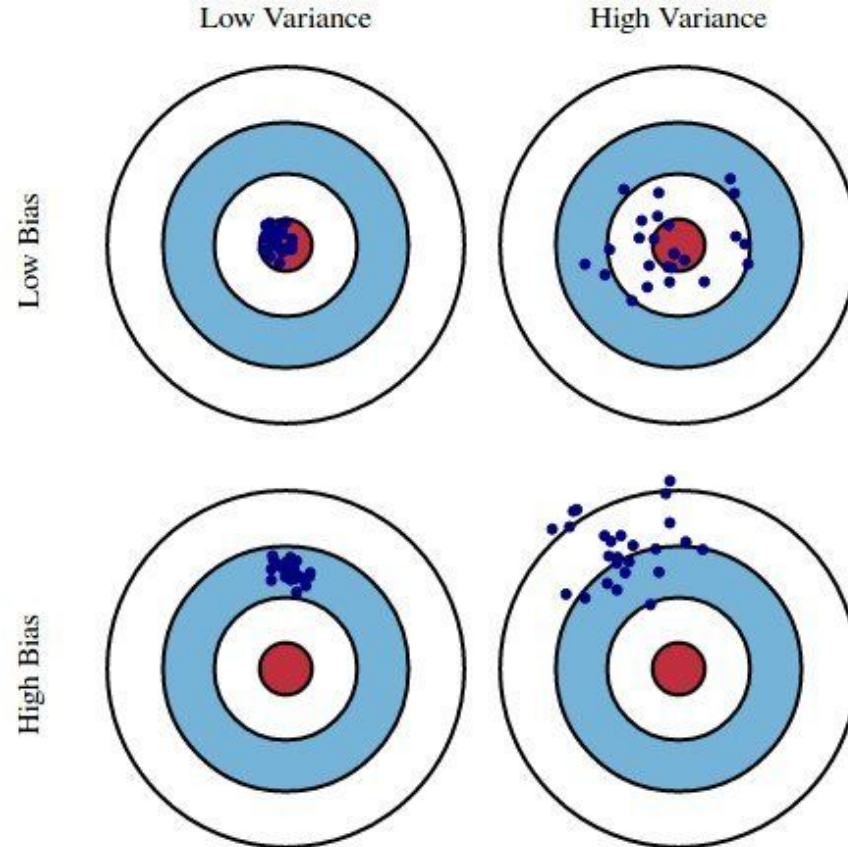
Analizando Viés vs Variância

- **O erro de viés** (bias) é considerado a diferença entre a previsão esperada (ou média) do nosso modelo e o valor correto que estamos tentando prever. Claro que você tem apenas um modelo, portanto, falar sobre valores de previsão esperados ou médios pode parecer um pouco estranho. No entanto, imagine que você possa repetir todo o processo de construção do modelo mais de uma vez: a cada vez que você reúne novos dados e executa uma nova análise, cria um novo modelo. Devido à aleatoriedade nos conjuntos de dados subjacentes, os modelos resultantes terão uma variedade de previsões. O erro de viés mede quão distantes, em geral, as previsões desses modelos estão do valor correto.

Analizando Viés vs Variância

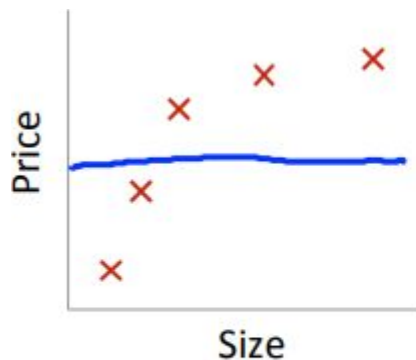
- **O erro devido à variância** é considerado como a variabilidade de uma previsão do modelo para um determinado ponto de dados. Novamente, imagine que você pode repetir todo o processo de construção do modelo várias vezes. A variância é o quanto as previsões para um determinado ponto variam entre as diferentes realizações do modelo.
- Essas ideias foram baseadas no artigo intitulado “[Accurately Measuring Model Prediction Error](#)” de Fortmann-Roe.

Analizando Viés vs Variância



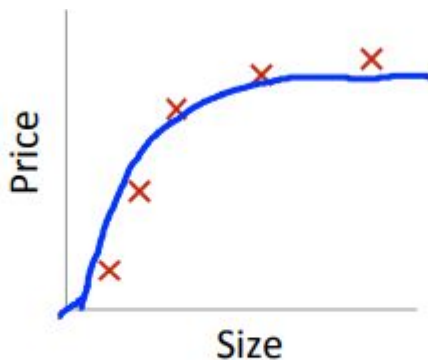
Regularização e Viés/Variância

- Como foi visto anteriormente, o efeito da regularização na hipótese do modelo tende a diminuir o sobreajuste nos dados (devido a um aumento do hiperparâmetro lambda).



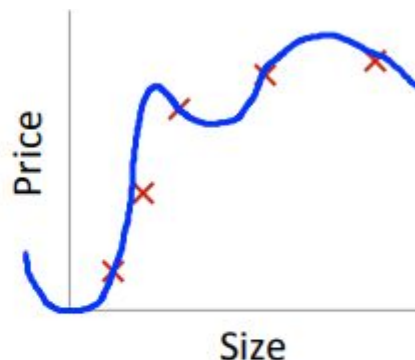
Large λ ←

► High bias (underfit)



Intermediate λ ←

"Just right"

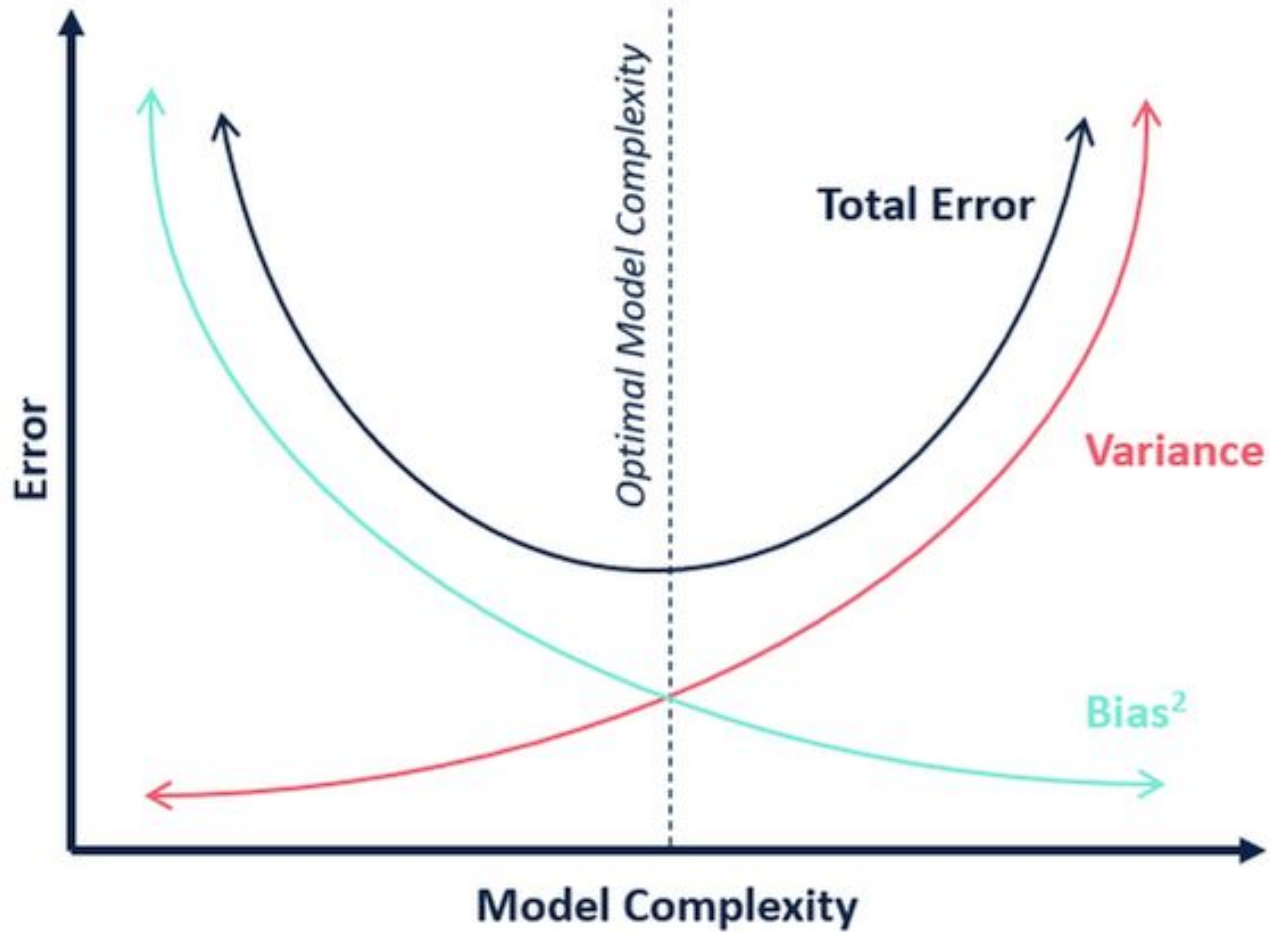


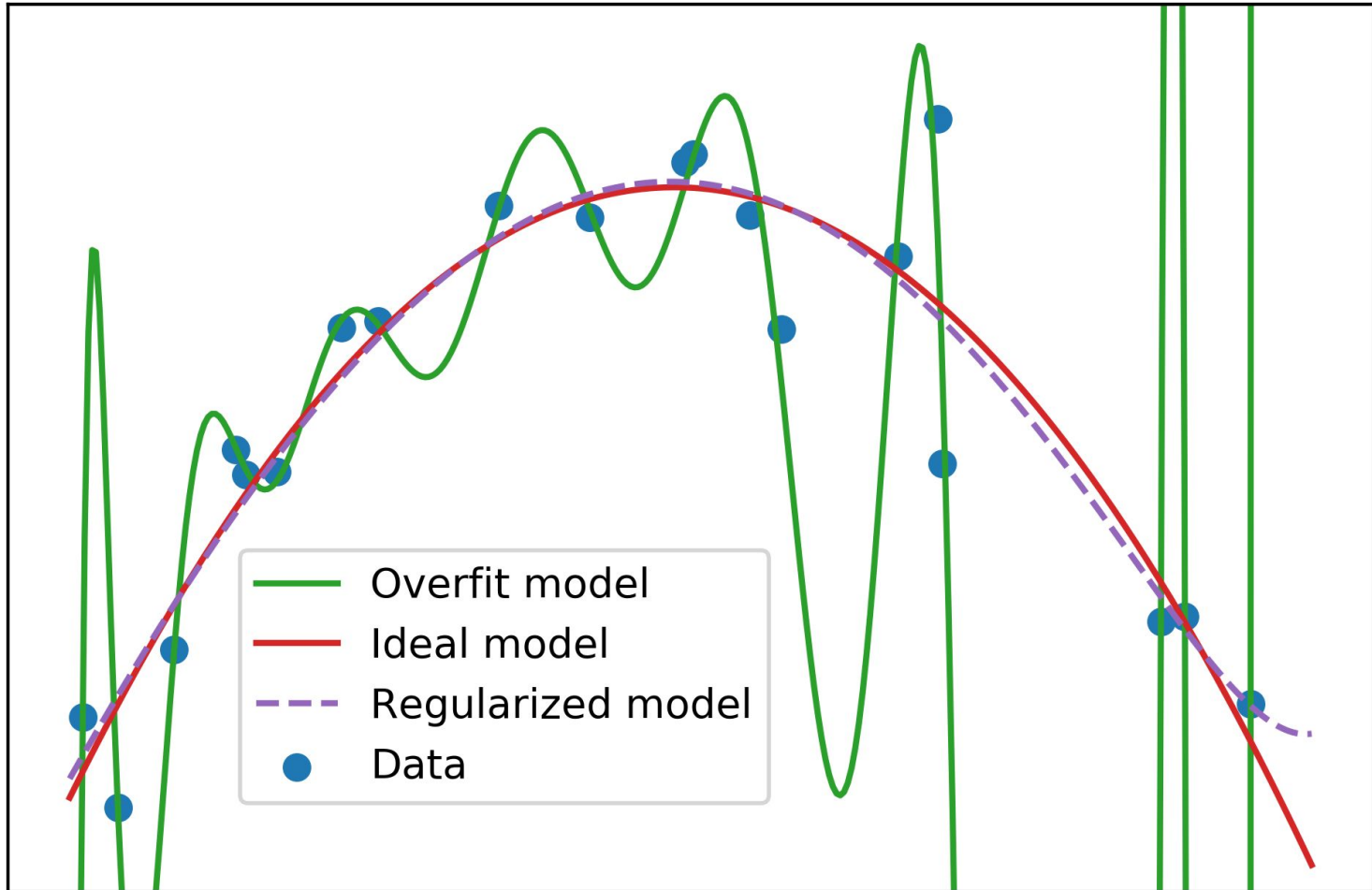
→ Small λ

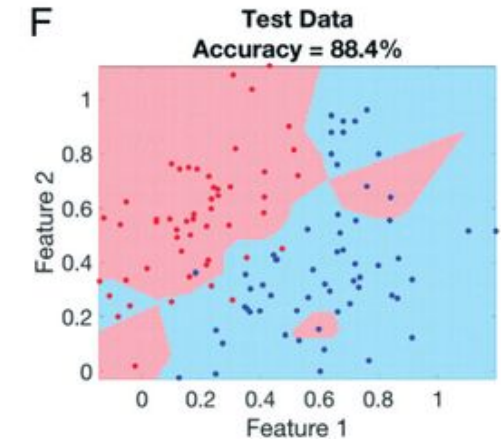
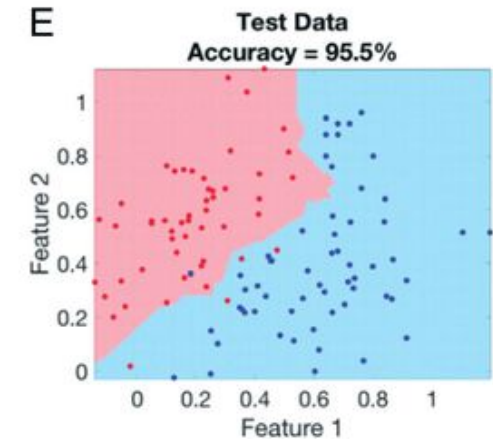
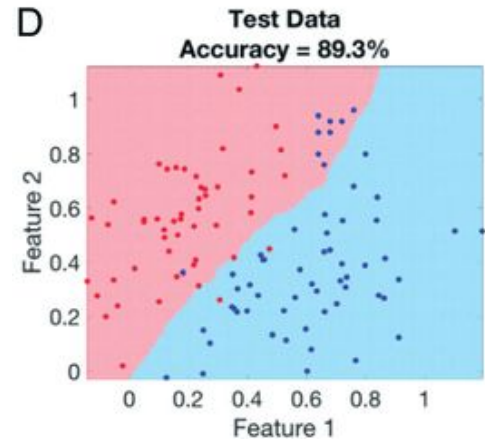
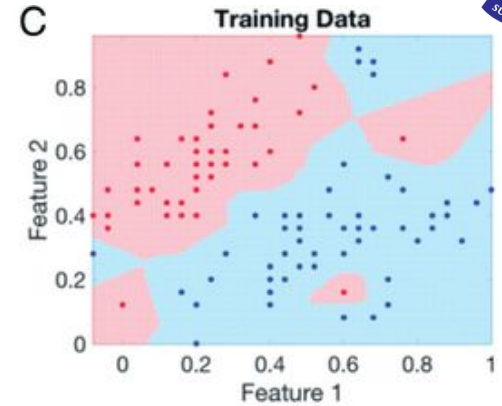
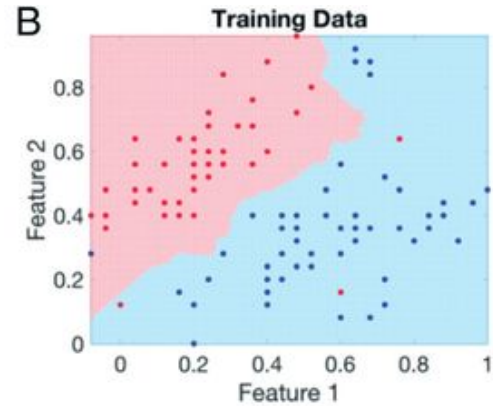
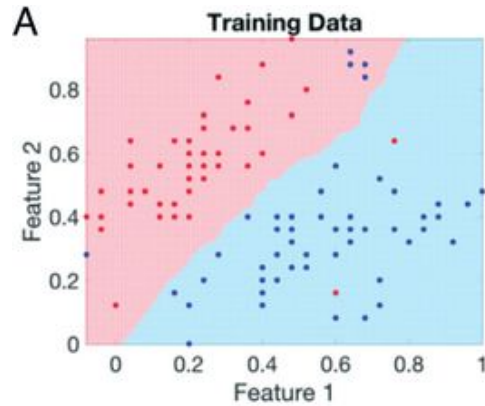
High variance (overfit)

Regularização e Viés/Variância

- Para escolher o modelo e o termo de regularização λ , precisamos:
 1. Crie uma lista de lambdas (ou seja, $\lambda \in \{0, 0.01, 0.02, 0.04, 0.08, 0.16, 0.32, 0.64, 1.28, 2.56, 5.12, 10.24\}$);
 2. Crie um conjunto de modelos com diferentes graus ou quaisquer outras variantes.
 3. Iterar através do λ se para cada λ passe por todos os modelos para aprender alguns Θ .
 4. Calcule o erro de validação cruzada usando o Θ aprendido (calculado com λ) no $J_{CV}(\Theta)$ **sem** regularização ou $\lambda = 0$.
 5. Selecione a melhor combinação que produz o menor erro no conjunto de validação cruzada.
 6. Usando a melhor combinação Θ e λ , aplique-a $J_{test}(\Theta)$ para ver se tem uma boa generalização do problema.





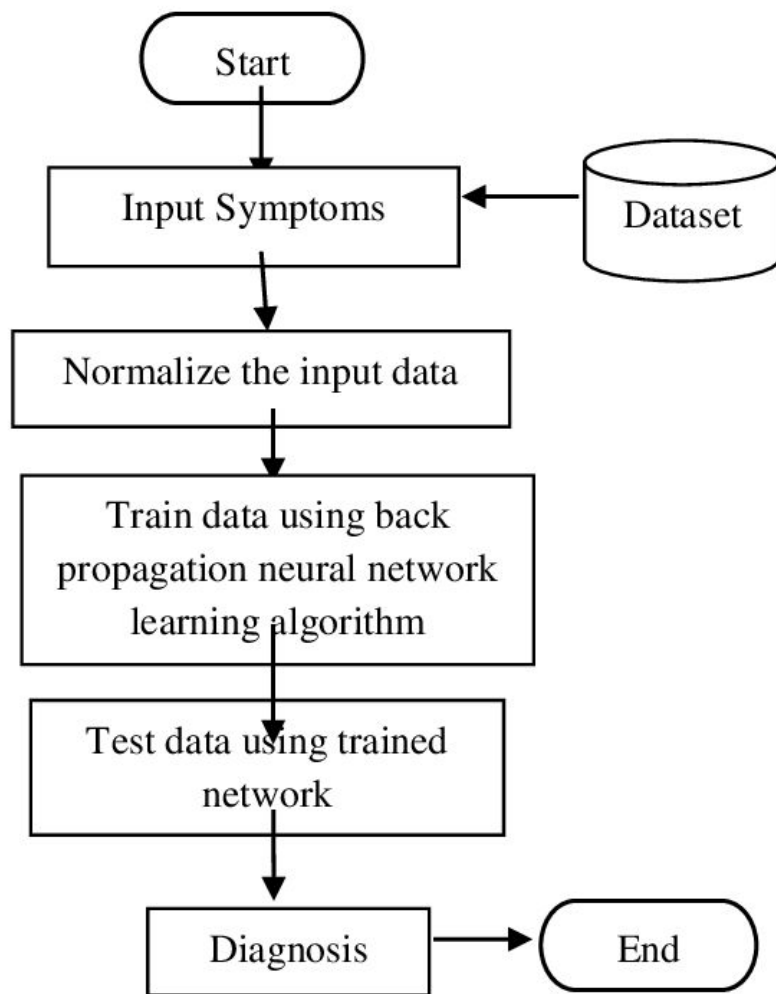


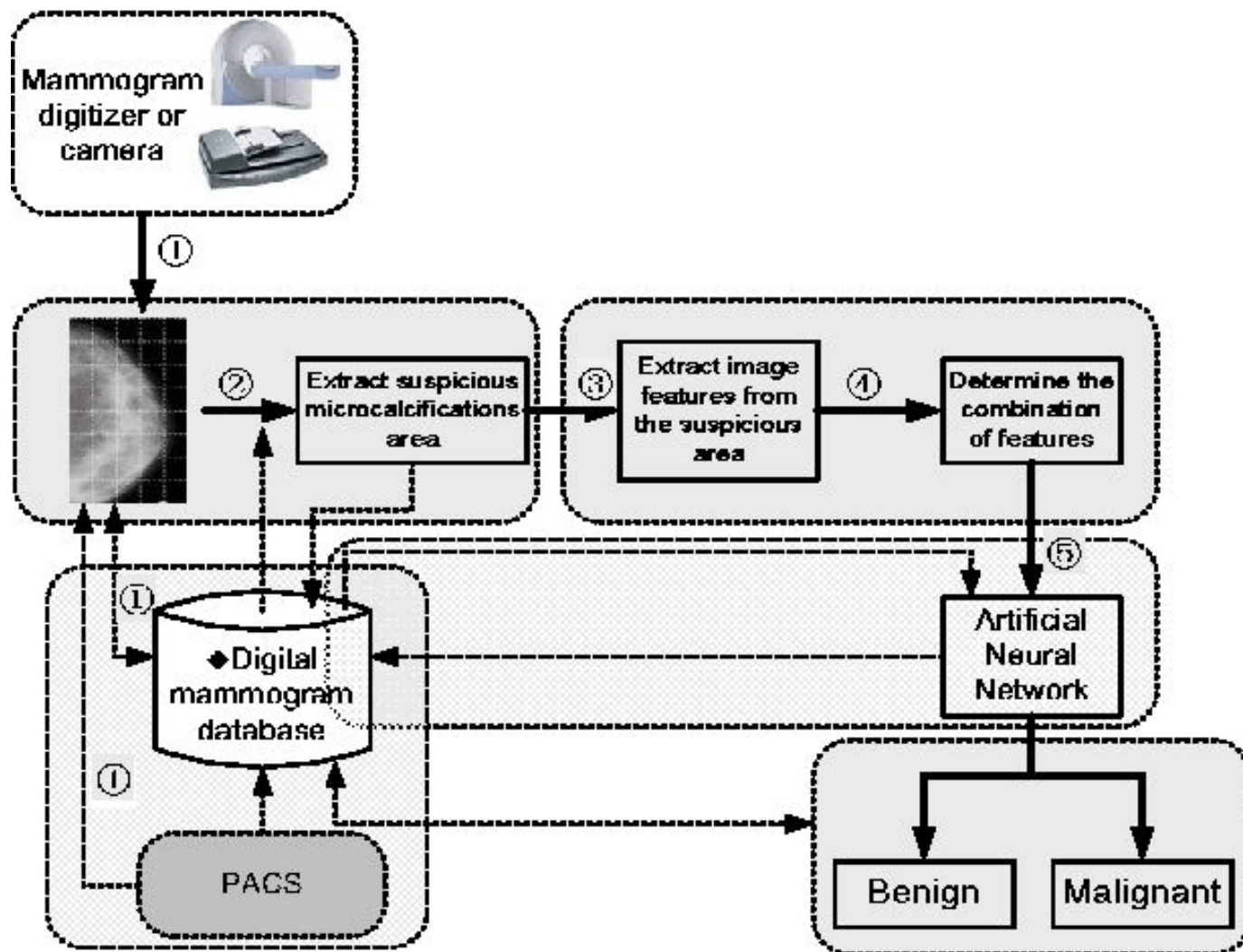
Revisando os passos

- **Obter mais exemplos de treinamento:** corrige alta variação
- **Tentar conjuntos menores de recursos:** corrige alta variação
- **Adicionando recursos:** corrige viés alto
- **Adicionar recursos polinomiais:** corrige viés alto
- **Diminuindo λ :** corrige o viés alto
- **Aumentar λ :** corrige a alta variância

Diagnóstico de Redes Neurais

- Uma rede neural com menos parâmetros tende a ser inadequada . Também é computacionalmente mais barato .
- Uma grande rede neural com mais parâmetros está sujeita a overfitting . Também é caro do ponto de vista computacional . Neste caso, você pode usar a regularização (aumentar λ) para resolver o sobreajuste.
- Usar uma única camada oculta é um bom padrão inicial. Você pode treinar sua rede neural em uma série de camadas ocultas usando seu conjunto de validação cruzada. Você pode então selecionar aquele com melhor desempenho.





Efeitos da Complexidade do Modelo

- **Polinômios de ordem inferior** (baixa complexidade do modelo) têm alto viés e baixa variância. Nesse caso, o modelo se ajusta de maneira pouco consistente.
- **Polinômios de ordem superior** (alta complexidade do modelo) ajusta-se aos dados de treinamento extremamente bem e aos dados de teste extremamente mal. Eles têm baixo viés nos dados de treinamento, mas variância muito alta.
- Na realidade, gostaríamos de escolher um modelo em algum ponto intermediário, que possa **generalizar bem, mas também se ajuste aos dados razoavelmente bem.**