

Feature Review

Machine Learning for High-Throughput Stress Phenotyping in Plants

Arti Singh,^{1,*} Baskar Ganapathysubramanian,²
Asheesh Kumar Singh,¹ and Soumik Sarkar²

Alunas: Mariana Silva Vianna e Natália Silva Morosini

Disciplina: LGN5831

PIRACICABA

2019

FENOTIPAGEM DE ESTRESSE VEGETAL NA AGRICULTURA

- **Stresses:**

- bióticos;
- abióticos.

- **Melhoramento genético de plantas:**

- incorporar genes resistentes e desenvolver cultivares resistentes.

- **Fenotipagem de plantas:**

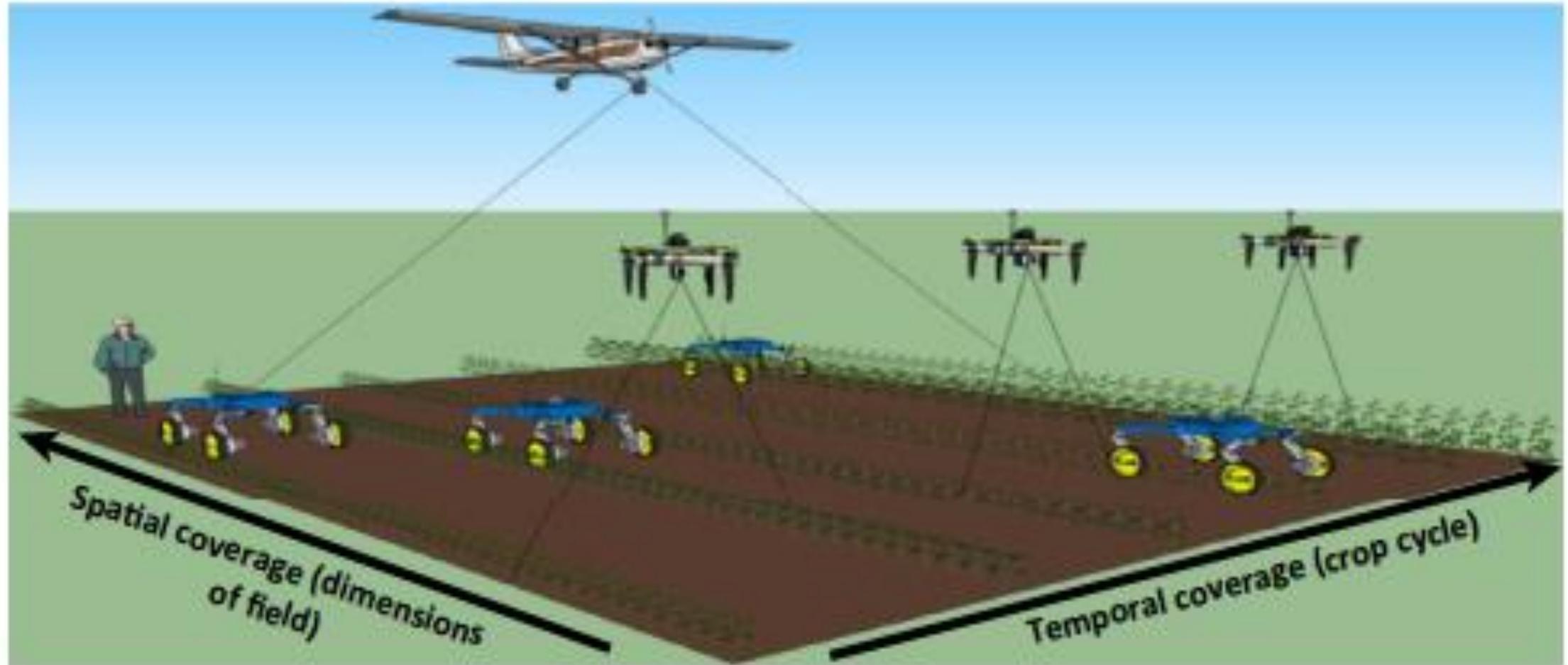
- várias características;
- grandes populações;
- ensaios replicados;
- múltiplos ambientes;
- *High-throughput phenotyping* (HTP) e *High-throughput stress phenotyping* (HTP).

DADOS DE FENOTIPAGEM E MACHINE LEARNING

- Plataformas em tempo real;
- Big data (problema);
- Impactos positivos na produção agrícola – promissor;
- Combinação de fatores (efeito holístico);
- HTSP habilitado para machine learning:

MACHINE LEARNING E HTP

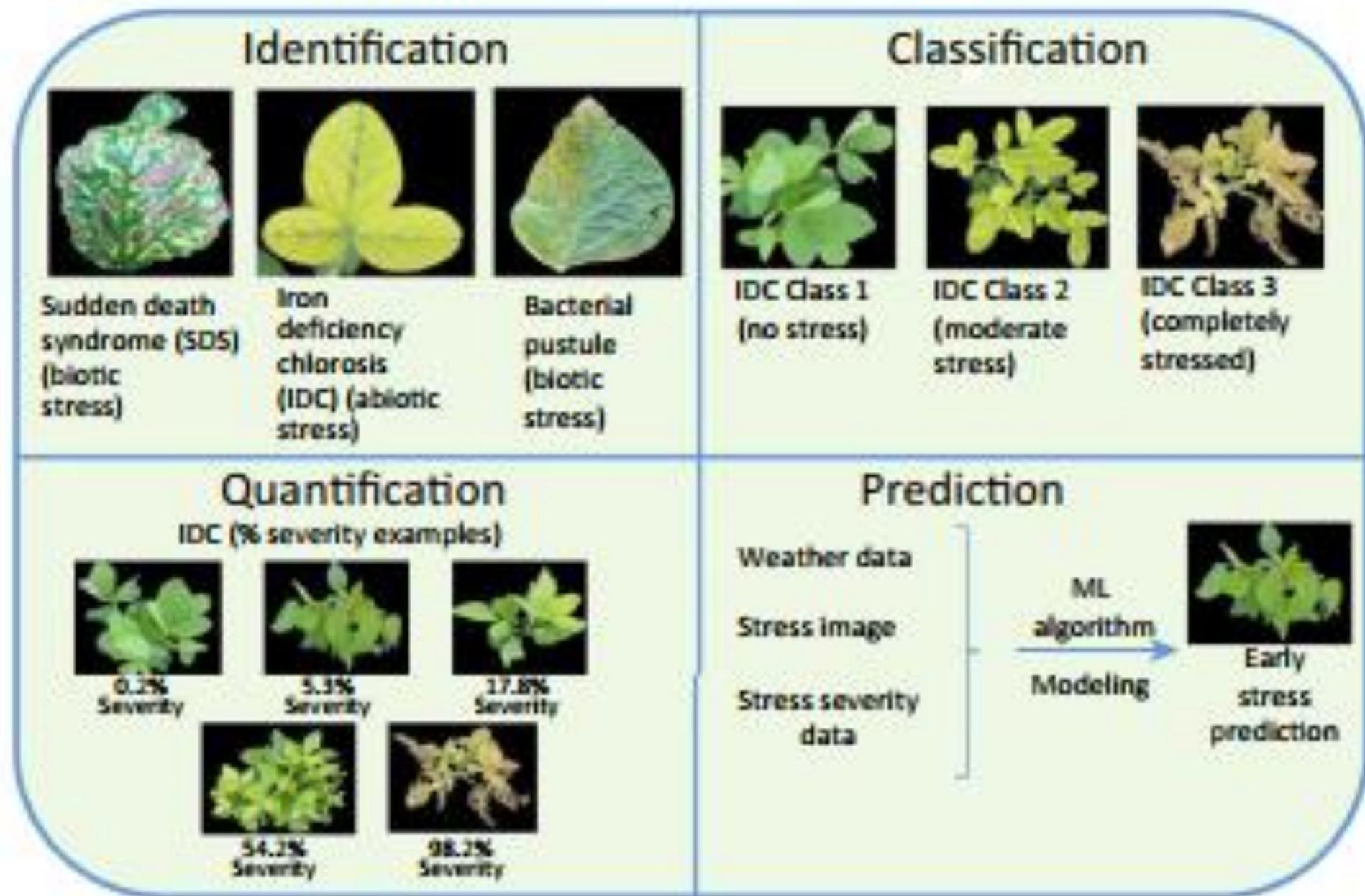
(A)



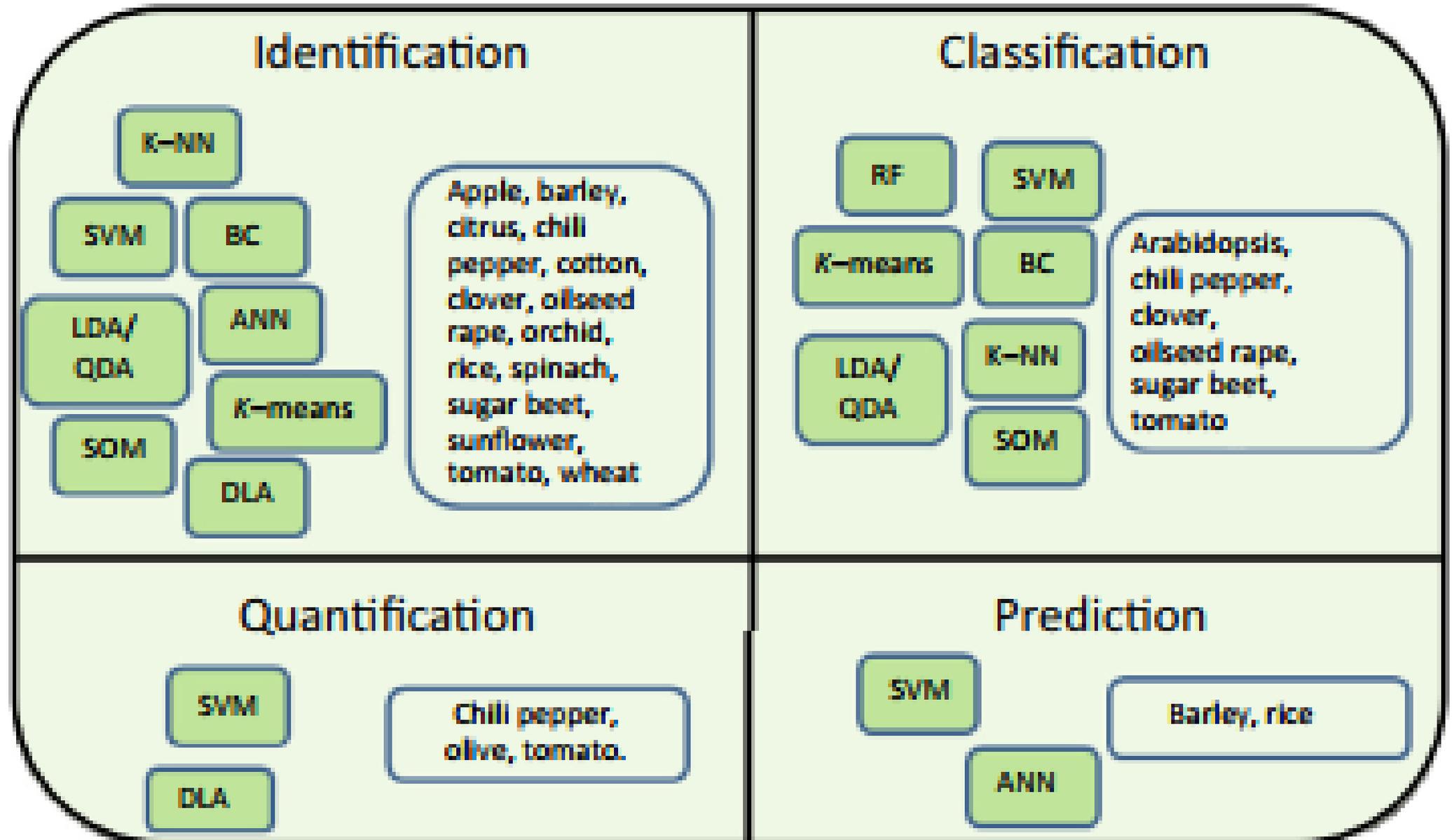
IIDENTIFICAÇÃO, CCLASSIFICAÇÃO, QQUANTIFICAÇÃO, PPREDIÇÃO (ICQP)

- O paradigma do ICQP caracteriza de forma abrangente o espectro de problemas na fenotipagem de estresse de plantas, permitindo assim o design e a implementação de ferramentas de ML apropriadas.

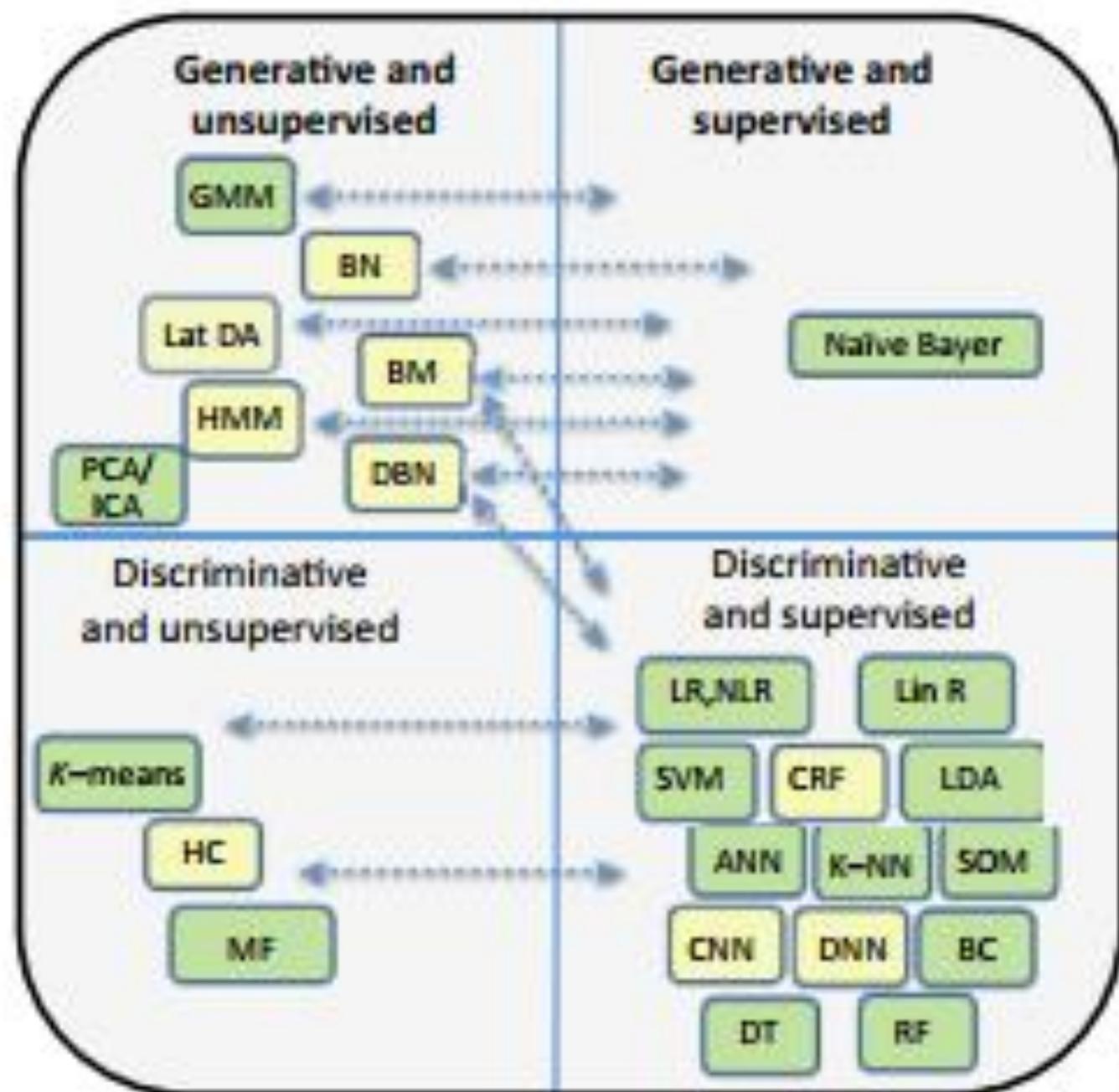
(B)



(C)



(D)



Key:



Used in plant stress phenotyping



No evidence of use in plant stress phenotyping



Arrows represent multiple variants of a method

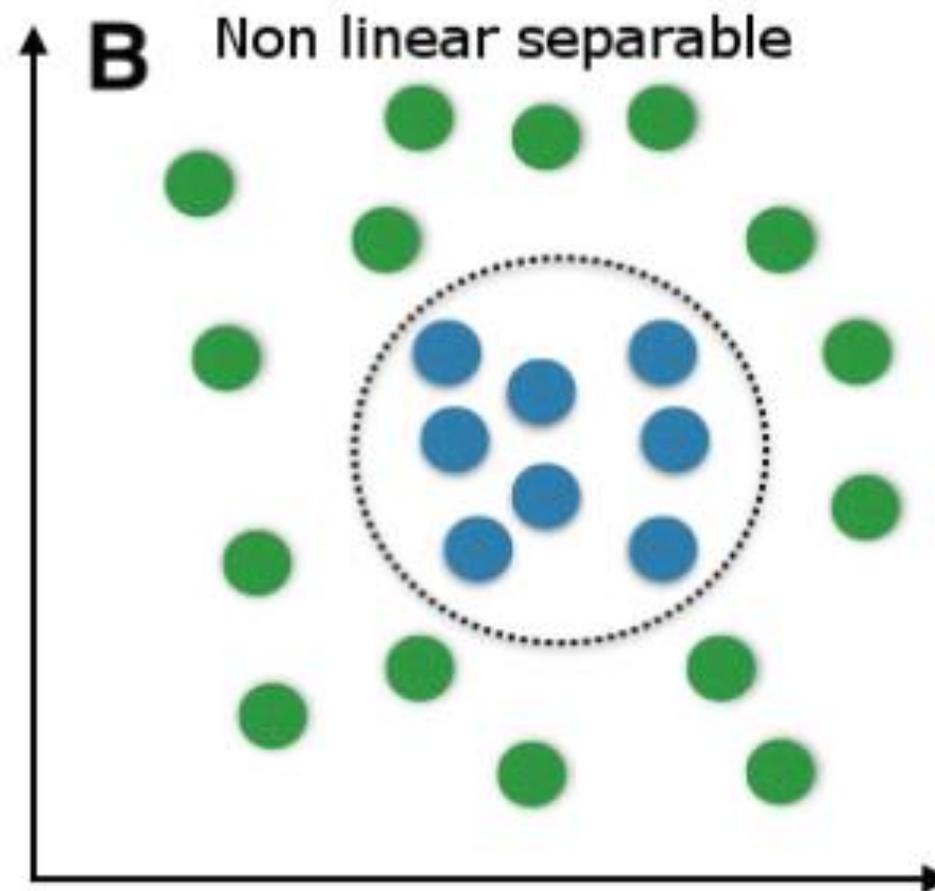
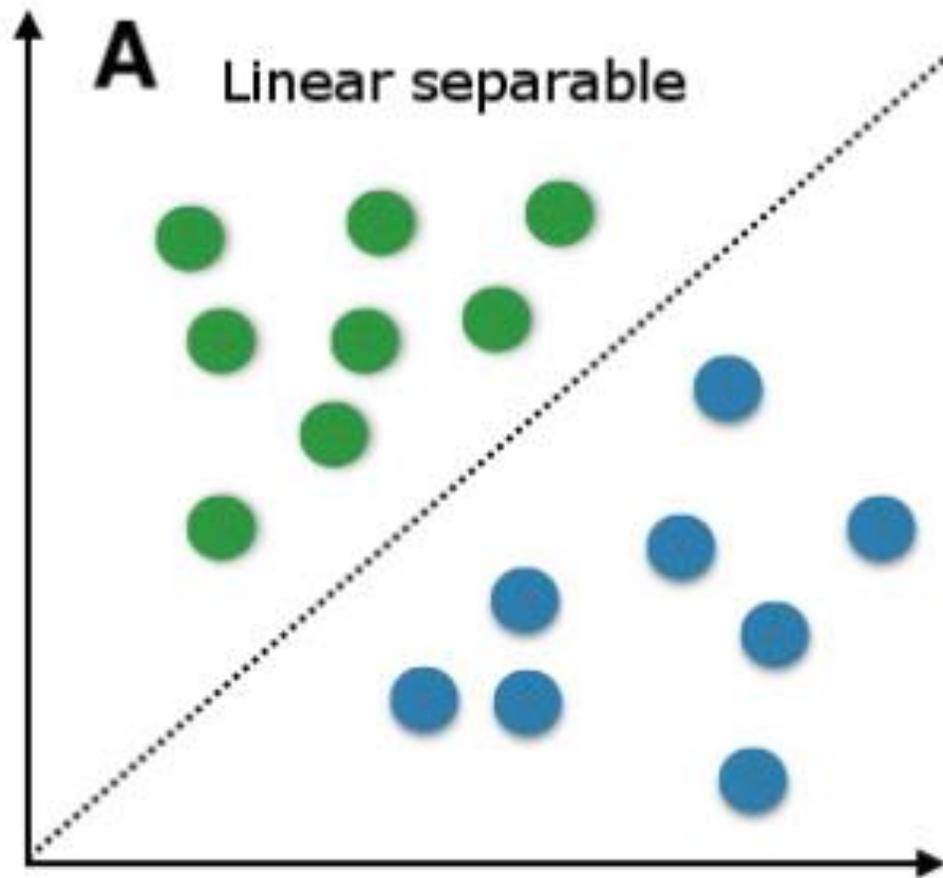
ABORDAGENS MACHINE LEARNING PARA ICQP

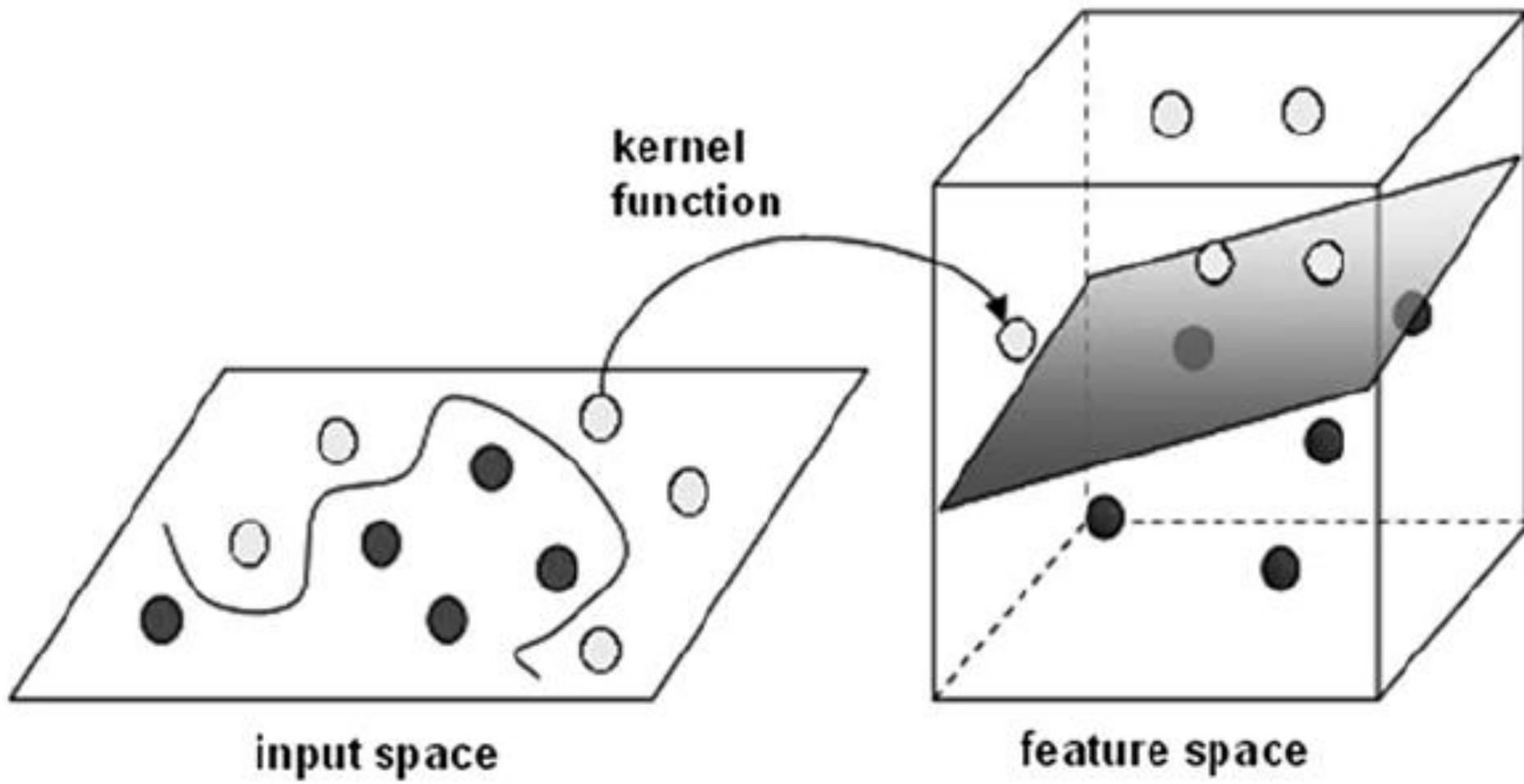
▪ Identificação do estresse:

- detecção de estresse específico;
- pré-processamento de dados de imagem é crucial;
- support vector machine (SVM) tem sido usado com sucesso em uma variedade de cenários para identificação de tensão em plantas.

MÉTODOSVM

- Citrus greening = Huanglongbing (HLB) – causada por bactéria:
 - estimativa de HLB – dados de espectroscopia de fluorescência (feito em folhas);
 - imagens pré-processadas -> features -> inputs no SVM.
- Duas plataformas de imagens aéreas foram comparadas para a identificação de doenças HLB em um pomar cítrico infectado:
 - imagens hiperespectrais;
 - imagens multibanda;
 - imagens pré-processadas -> features -> inputs no SVM.
- SVM (não-linear) com kernel funcionou melhor do que (linear) SVM, LDA e QDA.





ARTIFICIAL NEURAL NETS (ANN) E VARIANTES

- Processamento de imagem simples, seguido por ANN, foi usado para detectar e identificar três doenças das plântulas de orquídeas: podridão bacteriana, mancha marrom bacteriana e *Phytophthora* black rot.

ANÁLISES DISCRIMINANTES

- Dados de reflectância espectral foram registrados a partir de folhas cítricas saudáveis e infectadas com HLB;
- Análise discriminante quadrática (QDA);
- Análise discriminante linear (LDA) e k-vizinho mais próximo (k-NN);
- Modelagem independente de classe de analogia de classe (SIMCA)
- Na beterraba, a refletância do dossel hiperespectral foi utilizada para identificar sintomas de estresse - nematoide de cisto de beterraba e Rhizoctonia e podridão de raiz.

MODELO DE MISTURA GAUSSIANA

- O pré-processamento seguido pela aplicação de modelos de mistura gaussiana foi usado para identificar estresses bióticos e abióticos;
- Imagens de dossel foram tiradas em um campo de trigo estressado em níveis diferenciais de irrigação (estresse abiótico) e inoculados com vírus (estresse biótico).

K-CLUSTERING E VARIANTES

- O estresse hídrico e nutricional no nível da copa das culturas foi estudado usando imagens hiperespectrais que foram analisadas usando a maximização do volume simplex (SiVM), uma abordagem de clustering não supervisionada;
- O algoritmo SiVM foi encontrado para ser capaz de detectar o estresse hídrico usando imagens hiperespectrais 4 dias antes do aparecimento de sintomas visíveis ao olho humano.

REDUÇÃO DE DIMENSIONALIDADE E CLUSTER

- Algoritmo do fator Bayes com regressão por agregação de Dirichlet (DAR);
- Na cevada (*Hordeum vulgare* L.), o algoritmo DAR foi utilizado para a predição pré-sintomática do estresse hídrico em um estágio inicial, utilizando imagens hiperespectrais.

SOM E CLASSIFICADOR BAYES

- Abordagem supervisionada (classificador bayesiana) e não supervisionada (SOM) foram usadas para segmentar plantas de tomate doentes;
- Regiões doentes no dossel de tomate foram identificadas usando as imagens de tomate pré-processadas por um modelo SOM.

ABORDAGENS MACHINE LEARNING PARA ICQP

- Classificação do estresse:

- classifica o estresse em classes rotuladas com base em sintomas de estresse e assinaturas;
- estresse por seca pode ser classificado em: sem estresse, estresse moderado e estresse pesado.

- Exemplos:

- ANN;
- SVM;
- LDA , K-means e métodos acoplados.

ABORDAGENS MACHINE LEARNING PARA ICQP

- **Quantificação do estresse:**

- a severidade da ferrugem no trigo pode ser quantificada em uma escala de 0 a 100%. Os algoritmos de quantificação geralmente são melhor precedidos por um estágio de pré-processamento;
- usado para quantificar a severidade da doença de PHYVV (pimenta huasteco). vírus da veia amarela) em plantas de pimenta (*Capsicum annuum* L.).

- Exemplo:

- SVM.

ABORDAGENS MACHINE LEARNING PARA ICQP

- Predição do estresse:

- a previsão do estresse da planta em um estágio inicial, antes que seja visível ao olho humano, tem implicações substanciais para o controle oportuno e custo-efetivo do estresse. Há muito poucas atividades relatadas sobre o uso de ML para a previsão de tensões, tornando esta a próxima grande fronteira para os esforços de pesquisa.

CONCLUSÃO

- As perspectivas para as ferramentas de ML na agricultura são muito promissoras;
- Um ingrediente-chave para a aplicação em larga escala bem-sucedida do ML é a perfeita integração da análise de dados na coleta de dados;
- O ecossistema computacional abrirá enormes oportunidades para acelerar o melhoramento e resolver problemas fundamentais na genômica e nos fenômenos preditivos.