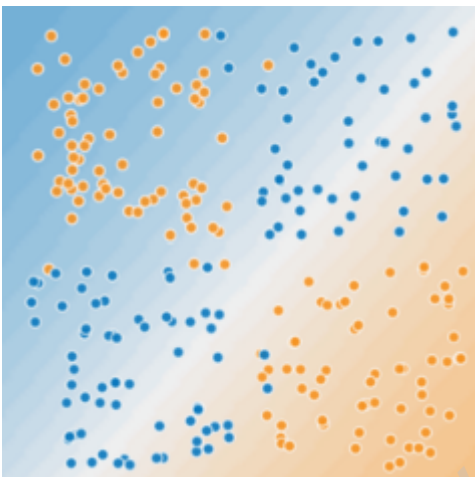


ML Concepts - Feature Crosses



**Figure 1



**Figure 2

- **Linear vs. Nonlinear Problems:**

- In Figure 1, a single line can neatly separate sick and healthy trees, indicating a linear problem.
- In Figure 2, no single straight line can effectively separate the two classes, indicating a nonlinear problem.

- **Feature Cross for Nonlinear Problems:**

- To address nonlinear problems, create a feature cross, which is a synthetic feature formed by multiplying two or more input features.
- The feature cross introduces nonlinearity into the feature space, allowing a linear algorithm to learn the relationships.

- **Example Feature Cross:**

- In the example, a feature cross named $A \times B$ is created by crossing features A and B .
- The linear formula incorporates this feature cross without changing the training process.

- **Types of Feature Crosses:**

- Feature crosses can take various forms, such as multiplying two feature values ($A \times B$), combining multiple features ($A \times B \times C \times D \times E$), or squaring a single feature ($A \times A$).

- **Efficient Training with Feature Crosses:**

- Linear models, thanks to stochastic gradient descent, can efficiently train on massive-scale datasets when supplemented with feature crosses.

Crossing One-Hot Vectors

- **Feature Crosses with One-Hot Vectors:**

- In practice, machine learning models often cross one-hot feature vectors, treating them as logical conjunctions.
- For example, crossing features like country (USA/France) and language (English/Spanish) results in binary features representing combinations like `country:usa` AND `language:spanish`.

- **Coarser Binning for Latitude and Longitude:**

- Using coarser bins for latitude and longitude leads to synthetic features representing broader ranges of combinations.

- **Enhancing Predictive Ability:**

- Feature crosses significantly enhance predictive ability. For instance, crossing behavior type and time of day in a dog owner satisfaction prediction can yield more meaningful insights than using each feature separately.

- **Efficient Scaling with Linear Learners:**

- Linear learners are effective with massive datasets, especially when combined with feature crosses. This strategy enables learning highly complex models efficiently. Neural networks provide an alternative approach for complex models.