Machine Learning
Candidate Elimination Algorithm

DR. BHARGAVI R

SCOPE

VIT CHENNAI

Introduction

- **Hypothesis Space**: The set of all possible hypotheses that can be formed using the given features and target variables.
- **Consistent Hypotheses**: Hypotheses that correctly classify all the training examples.
- **Version Space**: The subset of the hypothesis space that contains only the consistent hypotheses.
- Version space is the set of all hypotheses that are consistent with the observed training data.
- Version space can be visualized as a boundary between the most general hypothesis (which classifies everything as positive) and the most specific hypothesis (which classifies only the positive examples seen so far as positive).

Candidate Elimination Algorithm

Initialize:

G (General Boundary): The set of maximally general hypotheses.

S (Specific Boundary): The set of maximally specific hypotheses.

For each training example d:

If the example d is positive:

Remove from G any hypothesis inconsistent with d.

For each hypothesis s in S that is inconsistent with d

Remove s from S

Add to S all minimal generalizations h of s such that

h is consistent with d, and some member of G is more general than h

Remove from S any hypothesis that is more general than another hypothesis in S

If d is negative:

Remove from S any hypothesis inconsistent with d.

For each hypothesis g in G that is inconsistent with d

Remove g from G

Add to G all all minimal specializations h of g such that

h is consistent with d, and some member of S is more specific than h

Remove from G any hypothesis that is less general than another hypothesis in G

Candidate Elimination Algorithm Example

Example	*Sky*	*AirTemp*	*Humidity*	*Wind*	*Water*	*Forecast*	*EnjoySport*
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

Initialize : G0 = {< ?, ?, ?, ?, ?, ?>}

S0 = {<Φ, Φ, Φ, Φ, Φ, Φ>}

S0 <Φ, Φ, Φ, Φ, Φ, Φ>

S1 <Sunny, warm, Normal, Strong, warm, same>

S2 S3 <Sunny, warm, ?, Strong, warm, same>

S4 <Sunny, warm, ?, Strong, ?, ?>

< Sunny, warm, ?, ?, ?, ?> < Sunny, ?, ?, strong, ?, ?> < ?, warm, ?, strong, ?, ?>

G4 < Sunny, ?, ?, ?, ?, ?> < ?, warm, ?, ?, ?, ?>

< Sunny, ?, ?, ?, ?, ?> < ?, warm, ?, ?, ?, ?> < ?, ?, Normal, ?, ?, ?> < ?, ?, ?, ?, Cool, ?> < ?, ?, ?, ?, ?, same>

G3

G0 G1 G2 < ?, ?, ?, ?, ?, ?>