

دانشگاه آزاد اسلامی واحد تبریز



نام درس: یادگیری ماشین

بخش: یادگیری ماشین و مرتب سازی فضای فرضیه ها

نام استاد: دکتر مسعود کارگر

فهرست مطالب

- مرتب سازی فضای فرضیه
- مثالی از مفهوم یادگیری
- الگوریتم CANDIDATE-ELIMINATION

مرتب سازی فضای فرضیه ها

- یادگیری استقرایی یک شیوه یادگیری با ناظر است.
- یادگیری یک مفهوم کلی از روی چند مثال (جزء).
- به عبارت دیگر یادگیری تعمیم جزء به کل است.
- در این فصل هدف محدود کردن فضای عمومی فرضیه ها از یک حالت عمومی به یک فضای معین است که نهایتاً بتوان فرضیه سازگار با مثالها را یافت.

مثالی از مفهوم یادگیری

- هدف یادگیری مفهوم زیر از روی مثالها می باشد.
- " دوست من در چه روزهایی از ورزش آبی لذت می برد؟ "

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

- یک فرضیه از فضای فرضیه ها را با یک بردار ۶ عضوی نشان می دهیم.
 - ? یعنی هر مقدار قابل قبول است.
 - \emptyset یعنی هیچ مقداری قابل قبول نیست.
 - در غیراین صورت یک مقدار معین برای صفت معین می شود.
- یک فرضیه نمونه: $\langle ?, Cold, High, ?, ?, ? \rangle$

مثالی از مفهوم یادگیری

● Given:

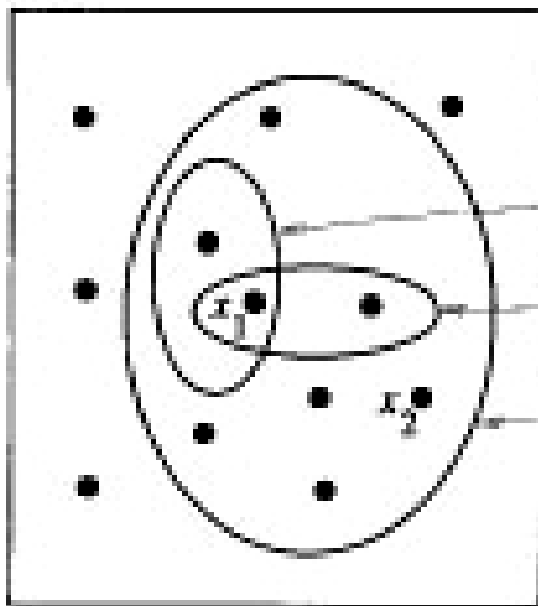
- Instances X : Possible days, each described by the attributes
 - *Sky* (with possible values *Sunny*, *Cloudy*, and *Rainy*),
 - *AirTemp* (with values *Warm* and *Cold*),
 - *Humidity* (with values *Normal* and *High*),
 - *Wind* (with values *Strong* and *Weak*),
 - *Water* (with values *Warm* and *Cool*), and
 - *Forecast* (with values *Same* and *Change*).
- Hypotheses H : Each hypothesis is described by a conjunction of constraints on the attributes *Sky*, *AirTemp*, *Humidity*, *Wind*, *Water*, and *Forecast*. The constraints may be "?" (any value is acceptable), "Ø" (no value is acceptable), or a specific value.
- Target concept c : $EnjoySport : X \rightarrow \{0, 1\}$
- Training examples D : Positive and negative examples of the target function (see Table 2.1).

● Determine:

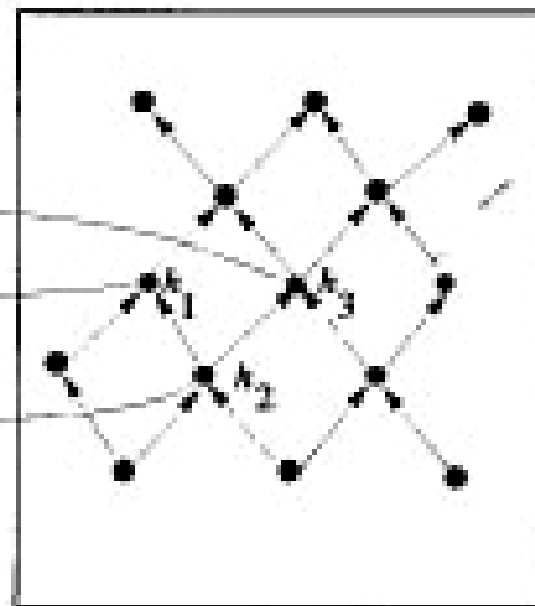
- A hypothesis h in H such that $h(x) = c(x)$ for all x in X .
-

مرتب‌سازی فرضیه‌ها از عمومی به معین

Instances X



Hypotheses H



Specific

General

$x_1 = \langle \text{Sunny, Warm, High, Strong, Cool, Some} \rangle$
 $x_2 = \langle \text{Sunny, Warm, High, Light, Warm, Some} \rangle$

$h_1 = \langle \text{Sunny, ?, ?, Strong, ?, ?} \rangle$
 $h_2 = \langle \text{Sunny, ?, ?, ?, ?, ?} \rangle$
 $h_3 = \langle \text{Sunny, ?, ?, ?, Cool, ?} \rangle$

الگوریتم CANDIDATE-ELIMINATION

- یک روش مرتب سازی فضای فرضیه هاست.
- مزیت این روش ارائه یک فرضیه سازگار (consistent) روی تمامی مثالهای آموزشی است.
- تعریف فضای نسخه Version Space: این فضا که با $VS_{H,D}$ نمایش داده می شود زیر مجموعه ای از فضای فرضیه های H را معین می کند که با همه مثالهای مجموعه D سازگار است.
- در حقیقت الگوریتم از هر دو سمت عمومی ترین و معین ترین فرضیه ها به سمت یافتن فضای نسخه حرکت می کند.

CANDIDATE-ELIMINATION الگوریتم

Initialize G to the set of maximally general hypotheses in H

Initialize S to the set of maximally specific hypotheses in H

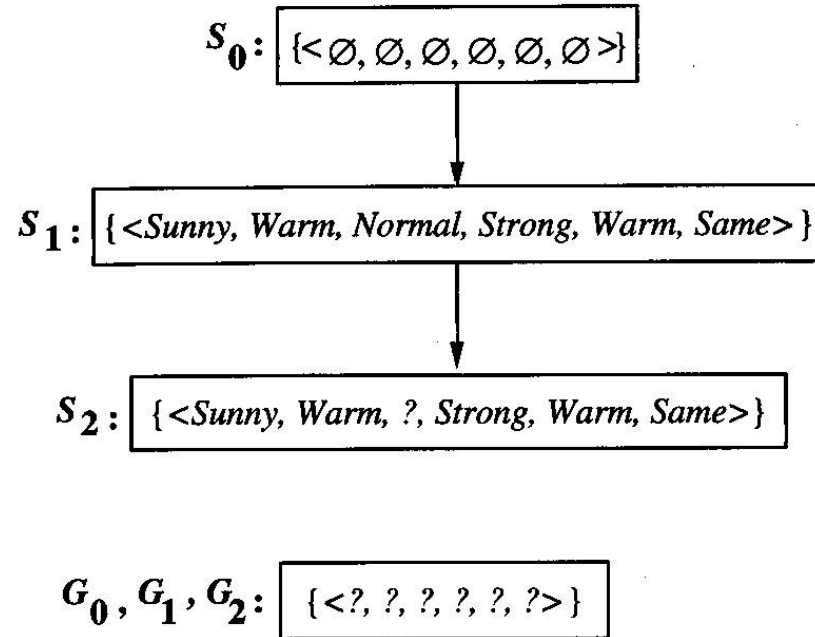
For each training example d , do

- If d is a positive example
 - Remove from G any hypothesis inconsistent with d
 - For each hypothesis s in S that is not consistent 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 a negative example
 - Remove from S any hypothesis inconsistent with d
 - For each hypothesis g in G that is not consistent with d
 - Remove g from G
 - Add to G 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

TABLE 2.5

CANDIDATE-ELIMINATION algorithm using version spaces. Notice the duality in how positive and negative examples influence S and G .

CANDIDATE-ELIMINATION الگوریتم



Training examples:

1. $\langle \text{Sunny}, \text{Warm}, \text{Normal}, \text{Strong}, \text{Warm}, \text{Same} \rangle, \text{Enjoy Sport} = \text{Yes}$
2. $\langle \text{Sunny}, \text{Warm}, \text{High}, \text{Strong}, \text{Warm}, \text{Same} \rangle, \text{Enjoy Sport} = \text{Yes}$

CANDIDATE-ELIMINATION الگوریتم

S_2, S_3 : { <Sunny, Warm, ?, Strong, Warm, Same> }

G_3 : { <Sunny, ?, ?, ?, ?, ?> <?, Warm, ?, ?, ?, ?> <?, ?, ?, ?, ?, Same> }

G_2 : { <?, ?, ?, ?, ?, ?> }

Training Example:

3. <Rainy, Cold, High, Strong, Warm, Change>, EnjoySport=No

CANDIDATE-ELIMINATION الگوریتم

S_3 : {<Sunny, Warm, ?, Strong, Warm, Same>}



S_4 : {<Sunny, Warm, ?, Strong, ?, ?>}

G_4 : {<Sunny, ?, ?, ?, ?, ?> <?, Warm, ?, ?, ?, ?>}

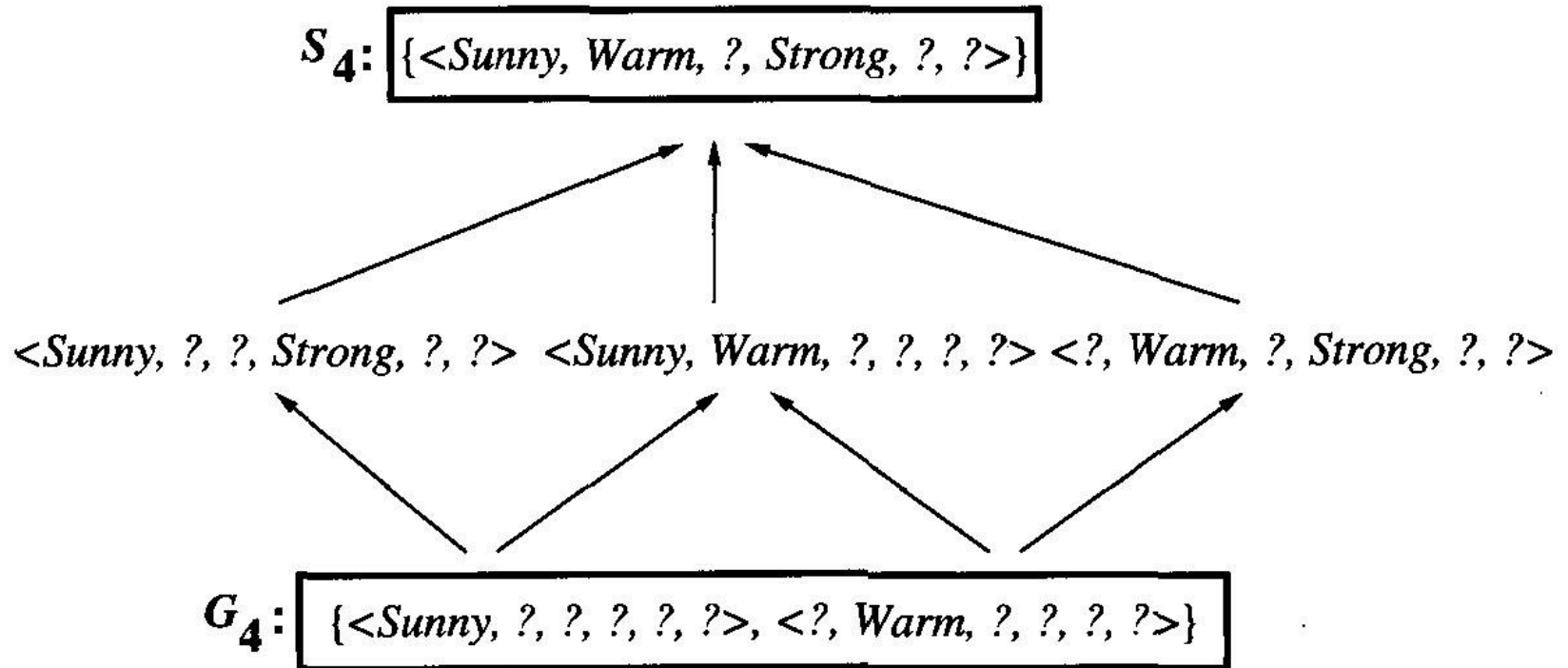


G_3 : {<Sunny, ?, ?, ?, ?, ?> <?, Warm, ?, ?, ?, ?> <?, ?, ?, ?, ?, Same>}

Training Example:

4. <Sunny, Warm, High, Strong, Cool, Change>, EnjoySport = Yes

الگوریتم CANDIDATE-ELIMINATION



فضای نسخه $VS_{H,D}$ محدود شده بین S_4 و G_4

قدردانی

- دکتر شیری:
- از دانشگاه صنعتی امیرکبیر