



Namatek
True Education



Decision Tree

www.namatek.com

درخت تصمیم

فهرست مطالب

۱. درخت تصمیم چیست؟
۲. کاربرد درخت تصمیم
۳. اجزای اصلی درخت تصمیم
۴. اندازه Decision Tree
۵. هرس کردن Decision Tree
۶. مزایای درخت تصمیم
۷. نرم افزارهای درخت تصمیم

داده کاوی موضوعی است که امروزه کاربرد زیادی در سازمان ها و شرکت ها به ویژه سیستم هایی که داده محور هستند، دارد و درخت تصمیم نیز یکی از ابزارهایی است که در داده کاوی مورد استفاده قرار می گیرد. این تکنیک زمانی به کار می آید که حجم داده ها بسیار زیاد است. در این مقاله قصد داریم شما را با مفهوم Decision Tree و کاربرد آن آشنا کنیم. با ما همراه باشید.

#1 درخت تصمیم چیست؟

درخت تصمیم (Decision Tree) ابزاری برای تصمیم گیری است تا اهداف، معیارها، گزینه ها و میزان سود و سرمایه گذاری را به تصمیم گیرنده نشان دهد. درخت تصمیم از نقاط یا به اصطلاح گره ها و شاخه ها تشکیل شده است که اولین گره آن نشان دهنده هدف تصمیم گیری است.

در ترسیم یک درخت تصمیم می توان از اشکال مختلفی برای نمایش اهداف و گزینه ها استفاده کرد. هر Decision Tree دربردارنده، انتخاب ها و پیشامدهای محتمل یا نتیجه این انتخاب ها است که در نهایت نتایج زنجیره، تصمیمات و این پیشامدهای احتمالی را نشان می دهد و تصمیم گیرندگان قادر هستند بر مبنای این نتایج تصمیمات خود را بهینه کنند و به نوعی از این نتایج برای پیش بینی وقایع آینده استفاده کنند.

زمانی که اطلاعات جدیدی به درخت تصمیم وارد می شود، این اطلاعات به صورت خودکار با اطلاعات قبلی سازمان دهی شده و سیستم کامل تر و با نظم بیشتری را ایجاد می کند که موجب می شود تصمیم گیرندگان تحلیل سیستماتیک بهتری را ارائه دهند. برای ایجاد یک Decision Tree باید موارد زیر در نظر گرفته شوند:

- شناسایی نقاط تصمیم و انتخاب های ممکن در هر سطح
- شناسایی احتمالات و بازه پیشامدها در هر سطح
- تخمین مقادیر عددی برای تحلیل احتمال نتایج عملکرد، هزینه ها و سود حاصل
- تحلیل ارزش انتخاب ها برای انتخاب یک مسیر

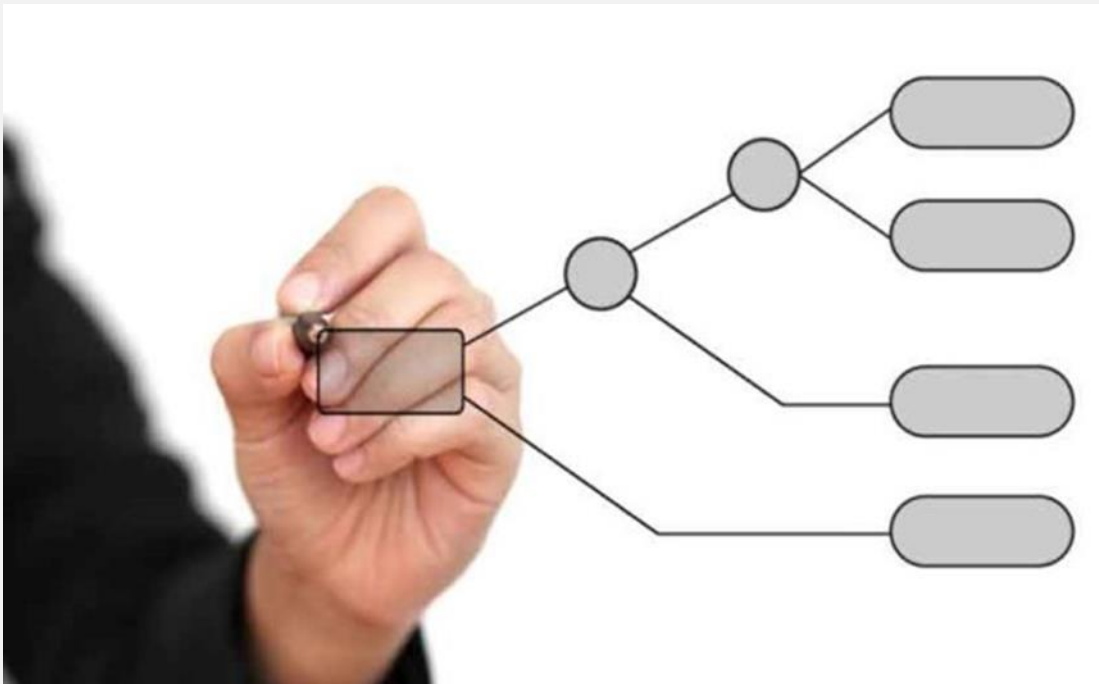
در یک Decision Tree همواره باید به این نکته توجه داشت که این درخت نتیجه قطعی و نهایی را به مدیران ارائه نمی دهد؛ بلکه مسیرهای بهینه را پیش پای آن ها قرار داده تا آن ها متناسب با سیاست های مورد نظر خود بهترین مسیر را انتخاب کنند و همچنین تحلیل بهتری از شرایط و پیش بینی بهتری از آینده داشته باشند.



#۲ کاربرد درخت تصمیم

همان طور که گفته شد، در درخت تصمیم علاوه بر قطعیت ها احتمالات هم دخالت دارند و بنابراین معیارها و گزینه ها علاوه بر حالت گسسته می توانند پیوسته هم باشند. این موضوع موجب می شود که در شرایط عدم قطعیت و پیش بینی آینده Decision Tree بسیار کاربردی باشد. درخت تصمیم بنا بر کاربردی که دارد طبقه بندی می شود. به طور مثال اگر

درخت برای مدل های رگرسیون استفاده شود، به عنوان درخت رگرسیون یا زمانی که برای کارهای دارای طبقه بندی استفاده شود، به عنوان درخت طبقه بندی شناخته می شود. اما در هر مدل از Decision Tree پیش بینی هایی که به دست می آید، بیانگر یک قانون و قواعد است که مطابق با آن تصمیم گیرنده تصمیمات خود را اتخاذ می کند.



#۳ اجزای اصلی درخت تصمیم

• برگ (Leaf Nodes)

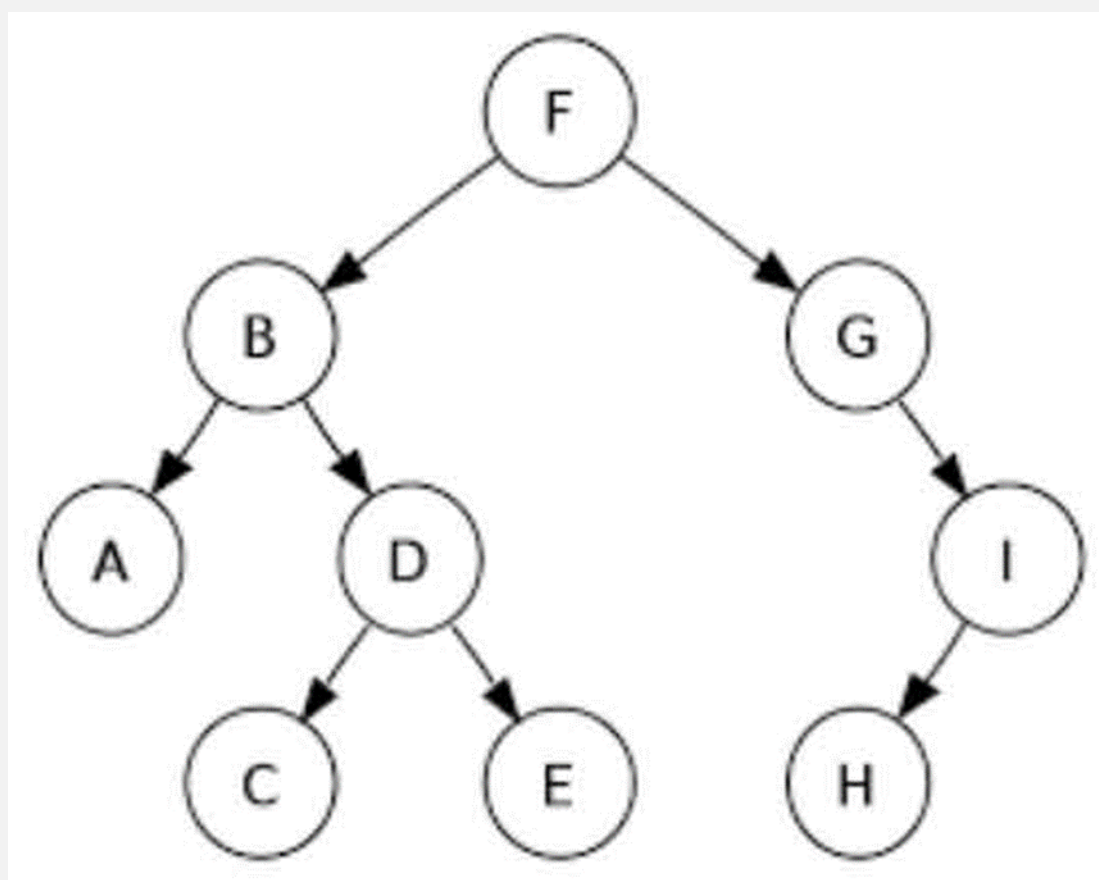
گره هایی که شروع و پایان تقسیم های متوالی هستند.

• ریشه (Root Node)

بیانگر گره آغازین در درخت تصمیم است.

• شاخه (Branches)

گره ها از طریق شاخه های به هم متصل می شوند که نشان دهنده روابط و مسیر بین معیارها و گزینه ها است.

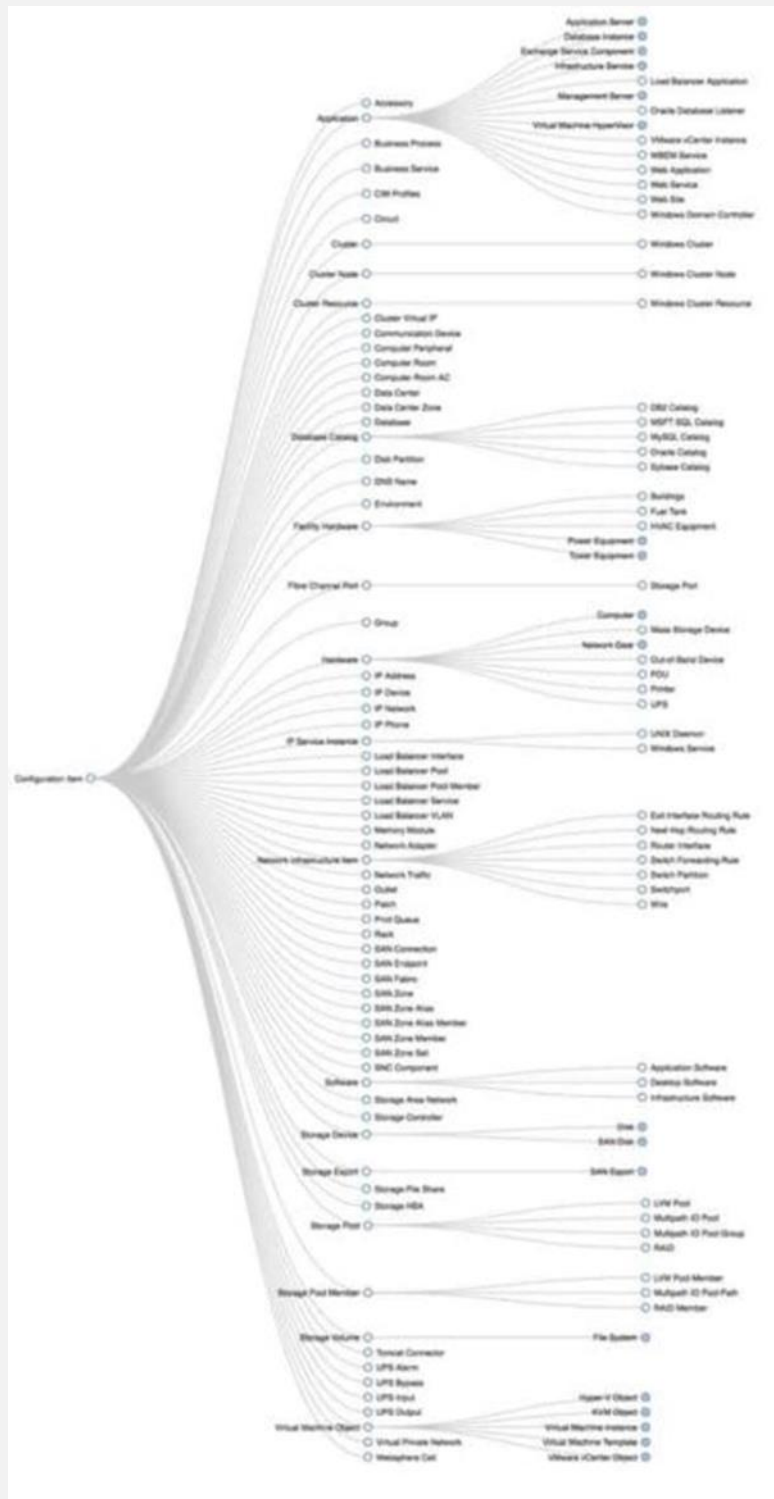


گره های درخت تصمیم به سه دسته تقسیم می شوند:

- گره های تصادفی (بیانگر احتمال وقوع نتایج است).
- گره های تصمیم گیری (تصمیمی که می توان گرفت را نشان می دهد).

- گره های پایانی (نشان دهنده نتیجه نهایی یک تصمیم گیری است.)

۴# اندازه Decision Tree



یک درخت تصمیم بنا بر هدف، معیار و گزینه هایی که در بر دارد و پیچیدگی که نشان می دهد، می تواند اندازه های مختلفی را داشته باشد.

میزان پیچیدگی یک درخت در اندازه آن تأثیر گذار است. این پیچیدگی می تواند در نتیجه تعداد کل گره ها، تعداد کل برگ ها، عمق درخت و تعداد مشخصه های به کار رفته باشد. البته پیچیدگی یک درخت را می توان به وسیله روش های مختلفی مانند معیار توقف و روش های به اصطلاح هرس کنترل کرد. در ادامه روش هرس کردن توضیح داده خواهد شد.

#۵ هرس کردن Decision Tree

با این کار بخش هایی از درخت تصمیم که باعث پیچیدگی آن شده است حذف می شود تا ناهنجاری های موجود در درخت برطرف و کنترل شوند. در نتیجه درخت ایجاد شده به راحتی قابل فهم می باشد و در تست های داده کاوی و تصمیم گیری بهتر از یک درخت پیچیده عمل می کند.

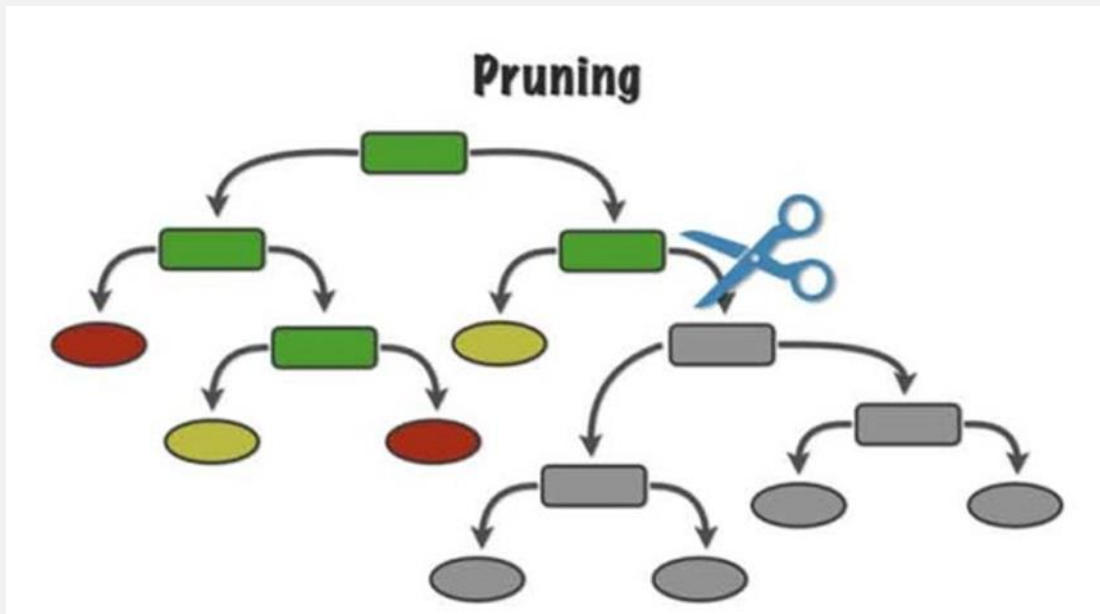
رویکردهای هرس درخت تصمیم

۱. پیش هرس

یک درخت به وسیله توقف های پشت سر هم در مراحل اول ساخت درخت ایجاد شده و بعد از هر توقف گره به برگ تبدیل می شود.

۲. هرس پسین

رایج تر از پیش هرس است. به این صورت که زیر درخت ها از یک درخت رشد یافته کامل حذف می شوند. در این حذف یک زیر درخت با یک برگ جایگزین شده و درخت تصمیم هرس می شود.



#۶ مزایای درخت تصمیم

همان طور که گفته شد ایجاد یک درخت تصمیم برای تصمیم گیری و داده کاوی به پیش بینی بهتر و تحلیل سیستماتیک شرایط به ویژه مواقع عدم قطعیت کمک می کند. علاوه بر موارد گفته شده مزایایی برای Decision Tree مطرح کرده اند که عبارتند از:

- یک ابزار بدیهی برای تصمیم گیری است که نیاز به توضیح به ذی نفعان و سایر تصمیم گیرندگان ندارد.
- می توان معیارهای عددی و اسمی یا کمی و کیفی را در آن وارد کرد.

- به طور خاص می توان تنها از درخت تصمیم برای نشان دادن طبقه بندی های مختلف استفاده کرد.
- داده هایی که احتمال خطا بودن آن ها زیاد است را نشان داده و تصمیم گیرنده را از وجود آن ها آگاه می کند.
- یک Decision Tree داده هایی که دارای مقدار پرت یا مفقود هستند را هم نشان می دهد تا در محاسبات در نظر گرفته نشوند.
- درخت های تصمیم روش های غیرپارامتری را هم مورد توجه قرار می دهند.
- برای تصمیم گیری مطابق با درخت تصمیم به محاسبات پیچیده نیاز نیست.
- معیارها با تأثیرگذاری بیشتر را نشان می دهند.
- در صورت لزوم امکان اضافه کردن گزینه های جدید به درخت وجود دارد.
- سازگار با دیگر ابزارهای تصمیم گیری است.



- تنها عیبی که برای درخت های تصمیم می توان گفت، این است که
- گاهی با ورود حجم بالایی از اطلاعات می توانند خیلی پیچیده شوند که در این صورت هم می توان با هرس کردن آن ها، پیچیدگی درخت تصمیم را بر طرف کرد.

#7 نرم افزارهای درخت تصمیم

استفاده از Decision Tree معمولا برای مواقعی کاربرد گسترده دارد که حجم داده ها بسیار زیاد باشد که حل دستی آن ها زمان گیر و هزینه بر است. به همین دلیل از نرم افزارهای مختلفی برای پیاده کردن Decision Tree استفاده می کنند.



نرم افزارهای زیر از الگوریتم درخت تصمیم برخوردار هستند:

- R
- SPSS Modeler
- Matlab
- SAS JMP
- Clementine
- Python

امروزه سازمان ها و شرکت ها از حجم داده بالایی برخوردار هستند که کار تصمیم گیری را برای آن ها مشکل کرده است. از طرفی به علت محدود بودن بودجه و زمان، نمی توانند تصمیم گیری های خود را به صورت دستی انجام دهند. از این رو استفاده از Decision Tree برای آن ها ضروری است. این موضوعی است که در آینده تمام شرکت ها و سازمان ها را شامل می شود؛ بنابراین لازم است تا با مفاهیم درخت تصمیم، نحوه کارکرد آن و نرم افزارهایی که برای محاسبات آن عرضه شده اند، آشنایی داشته تا بتوانند روند تصمیم گیری در شرکت را به راحتی و با نتایج صحیحی انجام دهند.