روش‌های کلی تشخیص نفوذ

Tuesday, 13 April 2021 19:20

شماره‌ی ۱ - منبع ۱ - ص. ۱ - ترجمه بدون تلخیص

روش تشخیص رفتار غیرعادی (Anomaly-based)

این روش سعی می‌کند تعیین کند که آیا می‌توان رفتار غیرعادی ایجاد شده را به عنوان یک نفوذ دانست یا خیر

روش تشخیص مبتنی بر امضا (Signature-based)

اما در این روش از الگوهای حملات انجام شده یا نقاط ضعف سیستم برای شناسایی نفوذ استفاده می‌شود.

K- نزدیک‌ترین همسایه (K-nearest neighbor)

19:46 Tuesday, 13 April 2021

شماره‌ی ۲ - منبع ۱ - ص. ۲ - ترجمه با تلخیص

روش K- نزدیکترین همسایه (k-NN) یکی از ساده‌ترین و سنتی‌ترین تکنیک‌های غیرپارامتری برای طبقه‌بندی نمونه‌ها است.
در این روش فاصله‌ی تقریبی بین نقاط مختلف بردارهای ورودی محاسبه می‌شود و سپس نقطه‌ی بدون برچسب به کلاس k-NN آن‌ها اضافه می‌شود.
در فرآیند ایجاد این طبقه بندی، k یک پارامتر مهم است و مقادیر مختلف آن باعث عملکردهای مختلف می شود. اگر k به طور قابل ملاحظه‌ای بزرگ باشد، همسایگانی که برای پیش‌بینی استفاده می‌کنند، زمان طبقه‌بندی زیادی دارند و بر دقت پیش‌بینی تأثیر می‌گذارند.
k-NN یادگیری مبتنی بر نمونه نامیده می‌شود و با رویکرد یادگیری استقرایی متفاوت است.

instance based learning

ماشین بردار پشتیبان (Support vector machines)

20:18 Tuesday, 13 April 2021

شماره‌ی ۳ - منبع ۱ - ص. ۲ - ترجمه با تلخیص

ماشین بردار پشتیبان (SVM) توسط Vapnik در سال ۱۹۹۸ ارایه شده است. SVM ابتدا بردار ورودی را در یک فضای با بُعد بالاتر ترسیم می کند و سپس بخش بهینه‌ای از آن را به دست می‌آورد. علاوه بر این، یک مرز تصمیم گیری، مانند همان محدوده‌ای که از فضای اصلی جدا شده، به جای کل نمونه‌های آموزشی توسط بردارهای پشتیبان تعیین می‌شود و بنابراین نسبت به نقاط دور از آن محدوده بسیار قوی است.
به طور خاص ، یک طبقه بندی SVM برای طبقه بندی به صورت باینری طراحی شده است. منظور از باینری این است که این روش، مجموعه‌ای از بردارهای آموزشی را که به دو کلاس مختلف تعلق دارند، جدا می‌کند.
توجه داشته باشید که **بردارهای پشتیبانی، نمونه‌های آموزشی نزدیک به مرز تصمیم‌گیری** هستند.
SVM همچنین یک پارامتر مشخص شده توسط کاربر به نام ضریب مجازات را فراهم می‌کند.
این پارامتر به کاربران این امکان را می‌دهد تا بین تعداد نمونه‌های طبقه بندی اشتباه و پهنای مرز تصمیم‌گیری معامله کنند.

penalty factor

شبکه‌های عصبی مصنوعی (Artificial neural networks)

21:05 Tuesday, 13 April 2021

شماره‌ی ۴ - منبع ۱ - ص. ۲ - ترجمه با تلخیص

شبکه‌ی عصبی یک واحد پردازش برای اطلاعات است که به تقلید از نورون‌های مغز انسان توسط Haykin در سال ۱۹۹۹ ابداع شده است.
پرسپترون چند لایه (MLP)، یکی از معماری‌های شبکه‌ی عصبی‌ست که به طور گسترده‌ای در بسیاری از مسائل تشخیص الگو استفاده می‌شود. یک شبکه‌ی MLP از یک لایه‌ی ورودی شامل مجموعه‌ای از گره‌های حسی به عنوان گره‌های ورودی، یک یا چند لایه‌ی مخفی از گره‌های محاسباتی و یک لایه‌ی خروجی از گره‌های محاسباتی تشکیل شده است.
هر اتصال داخلی با یک عدد به عنوان وزن آن اتصال همراه است که در مرحله آموزش تنظیم می‌شود.
برای آموزش MLP، معمولاً از الگوریتم یادگیری تولید متناوب استفاده می‌شود؛ به این شبکه‌ها، شبکه‌های عصبی انتشار مجدد نیز گفته می‌شود. در این شبکه‌ها، ابتدا وزن‌های تصادفی آموزش داده می‌شوند. سپس، الگوریتم وزن‌ها را تنظیم می‌کند تا برای هر چیز، یک واحد تعریف کند. این کار در به حداقل رساندن خطای طبقه بندی‌های غلط موثر است.

Multilayer perceptron (MLP)
backpropagation learning algorithm
backpropagation neural networks

نقشه‌های خود سازمان دهی شده (Self-organizing maps)

21:48 Tuesday, 13 April 2021

شماره‌ی ۵ - منبع ۱ - ص. ۲ - ترجمه بدون تلخیص

نقشه خود سازمان دهی شده (SOM) (Kohonen ، ۱۹۸۲) توسط الگوریتم یادگیری رقابتی بدون نظارت، آموزش داده می‌شود.
هدف SOM کاهش بُعد تجسم داده‌ها است. به این معنی که SOM بردارهای ورودی با ابعاد بالا را بر روی یک نقشه تصویری با ابعاد کم تجسم می‌کند که معمولاً این تصویربرای سادگی دو بُعدی است.
این الگوریتم معمولاً از یک لایه ورودی و لایه کوهونن تشکیل شده که به صورت آرایش دو بعدی نورون‌ها طراحی شده است و ورودی های n بعدی را در دو بعد ترسیم می کند.
لایه کوهونن وظیفه‌ی ایجاد ارتباط بین هر یک از بردارهای ورودی با یک خروجی نماینده را دارد.
شبکه، نزدیکترین گره به هر مورد آموزشی را پیدا می‌کند و سپس گره برنده را که نزدیک‌ترین نورون (یعنی نورون با حداقل فاصله) است، به عنوان گره آموزشی انتخاب می‌کند. یعنی SOM بردارهای ورودی مشابه را روی واحدهای خروجی یکسان یا مشابه روی یک نقشه دو بعدی ترسیم می‌کند. بنابراین، واحدهای خروجی خود را با یک نقشه‌ی مرتب سازماندهی می کند که و هم‌چنین واحدهای خروجی با وزن مشابه نیز پس از آموزش در همان نزدیکی قرار می گیرند.

Kohonen layer

درخت تصمیم (Decision trees)

22:21 Tuesday, 13 April 2021

شماره‌ی ۶ - منبع ۱ - ص. ۲ - ترجمه بدون تلخیص

درخت تصمیم یک نمونه را از طریق دنباله‌ای از تصمیمات طبقه بندی می‌کند، که در آن تصمیم فعلی به تصمیم گیری بعدی کمک می‌کند. چنین توالی تصمیماتی در یک ساختار درختی نشان داده می‌شود. طبقه‌بندی یک نمونه از گره ریشه به گره(ها)ی مناسب برگ منتهی می‌شود، جایی که هر گره برگ انتهایی نشان دهنده یک دسته‌بندی طبقه‌بندی شده‌است. ویژگی‌های نمونه‌ها به هر گره اختصاص می‌یابد و مقدار هر شاخه متناسب با صفات است (میچل ، ۱۹۹۷).

یک برنامه شناخته شده برای ساخت درختان تصمیم "طبقه بندی و بازگشت درخت" (CART) است (Breiman, Friedman, Olshen & Stone, 1984)

اگر به درخت تصمیم برچسب های گسسته یا نمادین کلاس‌بندی را بیفزاییم، درخت طبقه‌بندی نامیده می‌شود، در حالی که درخت تصمیم با دامنه‌ی مقادیر پیوسته یا عددی، درخت رگرسیون نامیده می‌شود.

classification tree
regression tree

شبکه‌های خلیج ساده (Naïve bayes networks)

22:25 Tuesday, 13 April 2021

شماره‌ی ۷ - منبع ۱ - ص. ۲ - ترجمه با تلخیص

موارد بسیاری وجود دارد که ما وابستگی‌های آماری یا روابط علت و معلولی بین متغیرهای سیستم را می‌دانیم. با این وجود، بیان دقیق روابط احتمالی میان این متغیرها ممکن است دشوار باشد. برای بهره برداری از این وابستگی‌های گاه به گاه بین متغیرهای تصادفی یک مسئله، می‌توان از یک مدل نمودار احتمالی به نام شبکه‌های خلیج ساده استفاده کرد.

الگوریتم‌های ژنتیکی (Genetic algorithms)

22:56 Tuesday, 13 April 2021

شماره‌ی ۸ - منبع ۱ - ص. ۳ - ترجمه با تلخیص

الگوریتم‌های ژنتیکی (GA) از کامپیوتر برای اجرای نظریه‌ی انتخاب طبیعی و تکامل استفاده می‌کنند. این الگوریتم توسط کوزا در سال ۱۹۹۲ پیشنهاد شده است.
الگوریتم با تولید تصادفی تعداد زیادی از برنامه‌های کاندیدا آغاز می‌شود. سپس از نوعی اندازه گیری تناسب اندام برای ارزیابی عملکرد هر فرد در یک جمعیت استفاده می‌شود. آن‌گاه تعداد زیادی تکرار انجام می‌شود تا برنامه‌های کم عملکرد با ترکیبات ژنتیکی برنامه‌های با عملکرد بالا جایگزین شوند. یعنی برنامه‌ای با اندازه گیری تناسب اندام کم حذف شده و برای تکرار بعدی کامپیوتر زنده نمی‌ماند.

منطق فازی (Fuzzy logic)

23:01 Tuesday, 13 April 2021

شماره‌ی ۹ - منبع ۱ - ص. ۳ - ترجمه با تلخیص

منطق فازی (یا نظریه مجموعه‌های فازی) مبتنی بر مفهوم پدیده فازی است که اغلب در دنیای واقعی رخ می‌دهد. نظریه مجموعه‌های فازی برای استدلال کردن، به مفهوم عضویت در مجموعه، مقادیر بین ۰ و ۱ را نسبت می‌دهد. یعنی در منطق فازی درجه حقیقت یک گزاره می‌تواند بین ۰ و ۱ باشد و محدود به دو مقدار حقیقت نیست (یعنی فقط درست و غلط). به عنوان مثال، "باران" یک پدیده طبیعی است و ممکن است بتواند شرایط محیطی را از عادی به اوضاع بحرانی تبدیل کند.

انواع طبقه‌بندی

23:16 Tuesday, 13 April 2021

شماره‌ی ۱۰ - منبع ۱ - ص. ۲ و ۳ - ترجمه با تلخیص

طبقه بندی منفرد

روش‌های که در این طبقه جای می‌گیرند تنها از یک الگوریتم برای حل مسئله تشخیص نفوذ استفاده می‌کنند. برای حل این مسائل از تکنیک‌های یادگیری ماشین (به عنوان مثال k-نزدیکترین همسایه، ماشین بردار پشتیبان، شبکه عصبی مصنوعی، درختان تصمیم، نقشه های خودسازماندهی شده و غیره) استفاده شده است.

طبقه بندی ترکیبی

ایده‌ی پشت طبقه بندی ترکیبی، ترکیب چندین تکنیک یادگیری ماشین است تا عملکرد سیستم به طور قابل توجهی بهبود یابد. به طور خاص، یک رویکرد ترکیبی معمولاً از دو جزو عملکردی تشکیل شده است. مورد اول داده‌های خام را به عنوان ورودی می‌گیرد و نتایج متوسطی را ایجاد می‌کند. سپس مورد دوم نتایج متوسط را به عنوان ورودی در نظر گرفته و نتایج نهایی را تولید می‌کند.