یادگیری ماشین

machine learning

انواع، تفاوت‌ها، کاربردها و الگوریتم‌های معروف

استاد : خانم دکتر الناز زعفرانی معطر

ارائه دهنده : میر امین موسوی نژاد

دسته‌بندی کلی یادگیری ماشین

یادگیری ماشین به سه دسته اصلی تقسیم می‌شود:

- یادگیری با نظارت Supervised Learning
- یادگیری بدون نظارت Unsupervised Learning
- یادگیری تقویتی Reinforcement Learning

- نکته:
- یادگیری عمیق (Deep Learning) یک روش پیاده‌سازی است که می‌تواند در هر سه دسته بالا استفاده شود، نه یک دسته مستقل.

یادگیری با نظارت Supervised Learning

- تعریف:

- الگوریتم با داده‌هایی آموزش می‌بیند که دارای برچسب Label هستند.
- هدف: یادگیری تابعی برای نگاشت ورودی به خروجی صحیح.
- انواع اصلی:

- طبقه‌بندی :Classification پیش‌بینی کلاس گسسته (مثلاً: اسپم/غیر اسپم).
- رگرسیون :Regression پیش‌بینی مقدار پیوسته (مثلاً: قیمت خانه).

- مثال‌های ساده
- ۱) تشخیص ایمیل اسپم Classification
- ورودی: متن ایمیل
- برچسب: اسپم هست یا نیست (بله/خیر)
- مدل با دیدن هزاران ایمیلِ برچسب‌دار یاد می‌گیرد که ایمیل جدید را هم اسپم تشخیص بدهد.

- ۲) پیش‌بینی قیمت خانه Regression
- ورودی: متراژ، تعداد اتاق، شهر...
- برچسب: قیمت واقعی (یک عدد پیوسته)
- مدل یاد می‌گیرد قیمت خانه‌های جدید را تقریب بزند.

چند الگوریتم معروف یادگیری با نظارت

- رگرسیون خطی (برای رگرسیون)
- رگرسیون لجستیک (برای طبقه‌بندی)
- KNN نزدیک‌ترین همسایه‌ها
- درخت تصمیم
- Random Forest
- SVM
- شبکه عصبی Neural Networks

تفاوت مهم یادگیری با نظارت و بدون نظارت با نظارت:

جواب درست داریم (برچسب وجود دارد)

- **بدون نظارت:**

جواب درست نداریم (مدل خودش الگو پیدا می‌کند)

یادگیری بدون نظارت یعنی چی؟

- در این نوع یادگیری، ما داده می‌دهیم ولی "جواب درست/برچسب" نداریم.
- یعنی:

- فقط ورودی داریم
- ولی نمی‌دانیم این داده‌ها دقیقاً متعلق به چه دسته‌ای هستند
- مدل خودش الگوها و ساختارهای پنهان داده را پیدا می‌کند

- مثال :
- چندین نفر را با قد و وزن داری، اما برچسب نداری که "این گروه ورزشکارهاست" یا "این گروه عادی‌هاست".
- مدل بدون نظارت بررسی می‌کند و خودش آن‌ها را بر اساس شباهت به چند گروه تقسیم می‌کند.

هدف‌های اصلی یادگیری بدون نظارت

- معمولاً ۳ هدف رایج داریم:

- ۱) خوشه‌بندی Clustering
- یعنی داده‌ها را به گروه‌های مشابه تقسیم کنیم.

- ۲) کاهش ابعاد Dimensionality Reduction
- وقتی داده‌ها خیلی ویژگی دارند (مثلاً صد عدد ویژگی)، مدل سخت می‌شود.

- ۳) قوانین وابستگی Association کشف ساختار
- مدل سعی می‌کند چیزهای تکرارشونده یا روابط پنهان را پیدا کند.

الگوریتم‌های معروف یادگیری بدون نظارت

- K-Means: محبوب‌ترین الگوریتم خوشه‌بندی.
- Hierarchical Clustering ایجاد ساختار درختی از خوشه‌ها.
- PCA تحلیل مؤلفه‌های اصلی: استانداردترین روش برای کاهش ابعاد.
- Apriori برای استخراج قوانین وابستگی (مثلاً سبد خرید).
- Autoencoders شبکه‌های عصبی برای کاهش ابعاد و بازنمایی داده‌ها.

یادگیری تقویتی Reinforcement Learning

- تعریف:

- یک عامل Agent در یک محیط Environment تعامل می‌کند.
- عامل با انجام اقدام، پاداش Reward یا جریمه Penalty دریافت می‌کند.
- هدف: یادگیری سیاست Policy برای حداکثر کردن پاداش تجمعی در طول زمان.

- تفاوت کلیدی:
- بازخورد فوری و برچسب‌دار ندارد؛ بلکه بازخورد تأخیری و مبتنی بر عملکرد است.

الگوریتم‌های معروف یادگیری تقویتی

- Q-Learning: یادگیری جدول ارزش برای جفت‌های وضعیت–اقدام.
- Deep Q-Networks (DQN): ترکیب Q-Learning با شبکه‌های عصبی عمیق.
- Policy Gradients: یادگیری مستقیم استراتژی عمل کردن.
- Actor-Critic: ترکیب دو روش بالا برای پایداری بیشتر.

ویژگی	یادگیری با نظارت	یادگیری بدون نظارت	یادگیری تقویتی
داده‌ها	دارای برچسب (جواب صحیح)	بدون برچسب	تعامل با محیط
هدف	پیش‌بینی یا طبقه‌بندی	کشف الگو یا گروه‌بندی	یادگیری بهترین تصمیم
بازخورد	فوری و دقیق (برچسب)	ندارد (خودکاشف)	پاداش/جریمه تأخیری
مثال	تشخیص ایمیل اسپم	بخش‌بندی مشتریان	بازی شطرنج با کامپیوتر

تفاوت دسته‌بندی و خوشه‌بندی

دسته‌بندی Classification:

نوع یادگیری: نظارت‌شده Supervised

برچسب‌ها: از قبل معلوم

هدف: تقریب تابع

$$f: X \rightarrow Y$$

که به هر ورودی x ، برچسب y را بدهد.

خروجی برای نمونه جدید: کلاس مشخص

خوشه‌بندی (Clustering):

- **نوع یادگیری: بدون نظارت (Unsupervised)**
- **هدف: پیدا کردن ساختار پنهان در داده (گروه‌های طبیعی)**
- **خروجی: نسبت‌دادن هر نمونه به یک خوشه** (Cluster ID، **مثل ۱،۲،۳...)**

دسته‌بندی Classification

یعنی دادن یک ورودی و انتخابِ یک کلاس درست از بین چند گزینه

کلاسترینگClustering

یعنی گروه‌بندی خودکار داده‌ها بدون برچسب بر اساس شباهت بینشان.

برای دسته‌بندی:

- Logistic Regression
- SVM (Support Vector Machine)
- Decision Tree / Random Forest
- Neural Networks / Deep Learning

برای خوشه‌بندی:

- K-means
- Hierarchical Clustering
- DBSCAN
- Gaussian Mixture Models (GMM)

RNN، LSTMو CNN

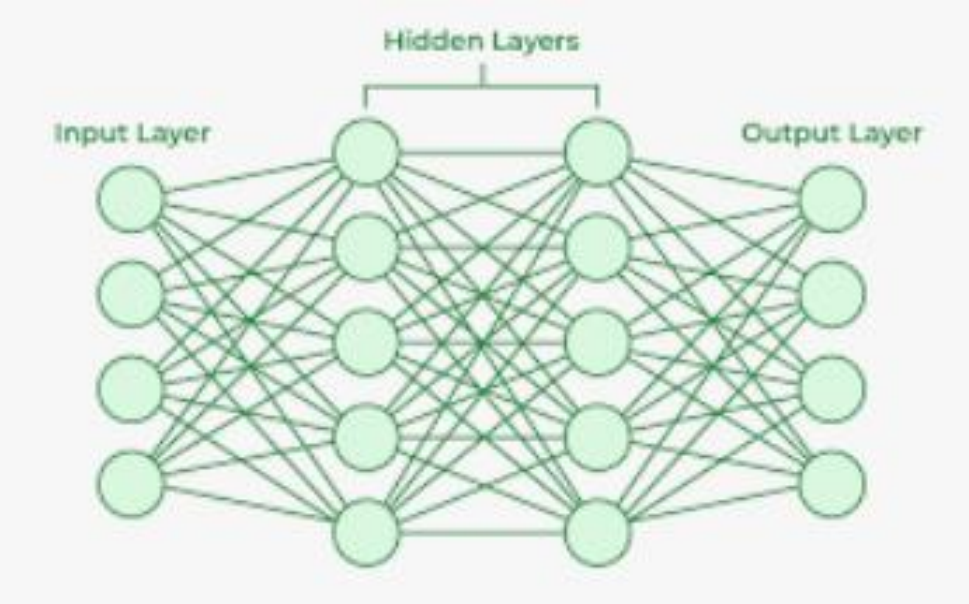
سه نوع از معروف‌ترین و پرکاربردترین شبکه‌های عصبی هستند که هرکدام برای نوع خاصی از مسائل طراحی شده‌اند.

شبکه عصبی کامپیوتر Artificial Neural Network یا ANN

یک مدل محاسباتی است که از ساختار و عملکرد شبکه‌های عصبی بیولوژیکی در مغز انسان الهام گرفته شده است. این شبکه‌ها برای یادگیری الگوها از داده‌ها طراحی شده‌اند و در وظایفی مانند تشخیص تصویر، پردازش زبان طبیعی، پیش‌بینی و تصمیم‌گیری کاربرد دارند.

ساختار کلی شبکه عصبی

یک شبکه عصبی از واحدهای کوچکی به نام "نورون" (یا گره) تشکیل شده است که در لایه‌های مختلف سازماندهی می‌شوند



لایه ورودی Input Layer

این لایه داده‌های خام را دریافت می‌کند. هر نورون در این لایه نشان‌دهنده یک ویژگی یا یک واحد از داده ورودی است.

لایه‌های پنهان Hidden Layers

این لایه‌ها بین لایه ورودی و خروجی قرار دارند. نورون‌های این لایه‌ها محاسبات پیچیده‌تری را انجام می‌دهند و الگوهای پنهان در داده‌ها را استخراج می‌کنند. تعداد لایه‌های پنهان و نورون‌های هر لایه به پیچیدگی مسئله بستگی دارد.

لایه خروجی Output Layer

این لایه نتیجه نهایی را تولید می‌کند، که می‌تواند یک پیش‌بینی، دسته‌بندی یا تصمیم باشد.

انواع شبکه‌های عصبی

شبکه‌های عصبی انواع مختلفی دارند که هر کدام برای وظایف خاصی بهینه‌سازی شده‌اند:

1. **شبکه‌های عصبی پرسپترون چندلایه**
2. **شبکه‌های عصبی کانولوشنی CNN**
3. **شبکه‌های عصبی بازگشتی RNN –Recurrent Neural Networks**
4. **شبکه‌های عصبی ترنسفورمر Transformer Networks**
5. **شبکه‌های مولد تخاصمی Generative Adversarial Networks - GAN**
6. **شبکه‌های عصبی خودرمزگذار Autoencoders**

شبکه RNN

- شبکه‌ای که داده‌ها را یکی‌یکی (پشت سر هم) پردازش می‌کند.
- و مهم‌تر از همه، در هر مرحله، یک «حافظه» از آنچه قبلاً دیده را به مرحله‌ی بعدی منتقل می‌کند.

این «حافظه» باعث می‌شود شبکه‌ی عصبی بتواند الگوهای زمانی یا ترتیبی را یاد بگیرد.

ایده کلی:

حافظه‌ی جدید = (اطلاعات ورودی جدید + اطلاعات گذشته) با هم مخلوط شود و یک "غیرخطی‌سازی" میدهد.

ورودیِ جدید رو با وزن‌ها پردازش کن

یک عدد ثابت برای تنظیم

$$h_t = \underline{f}\left(W_{xh}x_t + W_{hh}h_{t-1} + b_h\right)$$

یک تابع غیرخطی

حافظه‌ی قبلی رو هم با وزن‌ها پردازش کن

تابع غیرخطی چیست؟

تابع غیرخطی تابعی است که رابطه بین ورودی و خروجی در آن «مستقیم نیست».

LSTM

- ساختار LSTM (Long Short-Term Memory)

LSTMنوعی RNNهوشمندتر است که مخصوصاً طراحی شده تابتواند وابستگی‌های طولانی‌مدت را بهتر حفظ کند

مشکل گرادیان ناپدید شدن را کاهش دهد

ایده‌ی اصلی LSTMبه‌جای یک حافظه‌ی ساده ht، دو چیز دارد:

حالت مخفی: ht

حالت سلول (حافظه‌ی طولانی‌مدت): ct

و داخل هر سلول LSTM، چند «دروازه» Gate داریم که تصمیم می‌گیرند:

چه اطلاعاتی را فراموش کنیم

چه اطلاعاتی را اضافه/به‌روزرسانی کنیم

چه چیزهایی را به عنوان خروجی بدهیم

اجزای اصلی LSTMدر زمان t

ورودی‌ها: xt

ورودی فعلی : ht−1

حالت مخفی قبلی : ct−1حافظه‌ی قبلی

دروازه فراموشی (Forget Gate)
تصمیم می‌گیرد چه بخش‌هایی از حافظه‌ی قبلی را دور بریزد

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f)$$

خروجی f_t بین ۰ و ۱ است.
هر مؤلفه نزدیک ۰ ⇒ آن بخش از حافظه را فراموش میکند.
نزدیک ۱ ⇒ نگه دار.

دروازه ورودی (Input Gate)
تصمیم می‌گیرد چه اطلاعات جدیدی را ذخیره کند

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i)$$

و کاندیدای حافظه‌ی جدید:

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c)$$

به‌روزرسانی حافظه (Cell State Update)
ترکیب حافظه‌ی قبلی و جدید

دروازه خروجی (Output Gate)
تصمیم می‌گیرد از حافظه چه چیزی به عنوان خروجی بدهد

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o)$$

آپدیت حافظه و خروجی

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$

بخش اول: نگه‌داری/فراموشی حافظه قبلی

تابع هیپربولیک

$$h_t = o_t \odot \tanh(c_t)$$

بخش دوم: اضافه شدن اطلاعات جدید

اول حافظه را با tanh محدود می‌کنیم بعد با دروازه خروجی تعیین می‌کنیم چه مقدار از آن را خروجی کنیم

کاربردهای اصلی LSTM

۱) پردازش متن (NLP)

- تحلیل احساسات (مثبت/منفی) با توجه به کل جمله
- ترجمه ماشینی
- برچسب‌زنی متن (مثل NER)
- مدل زبانی و پیش‌بینی کلمه بعدی

چرا LSTM؟ چون روابط بین کلمات در فاصله‌های طولانی را بهتر از RNN ساده نگه می‌دارد.

۲) سری‌زمانی و پیش‌بینی

- پیش‌بینی قیمت/نرخ‌ها (بورس، ارز، سهام)
- پیش‌بینی مصرف برق
- پیش‌بینی ترافیک یا ازدحام
- پیش‌بینی تقاضای فروش

۳) تشخیص و تحلیل صدا

- تشخیص گفتار
- تحلیل سیگنال گفتار در بازه‌های زمانی
- تشخیص گوینده یا ویژگی‌های آکوستیکی

۴) ویدئو و تحلیل حرکت

- پیش‌بینی قاب بعدی یا تشخیص رویداد در ویدئو
- ردیابی حرکت/اکشن recognition (مثلاً تشخیص "راه رفتن" یا "دویدن")

۵) کنترل و تصمیم‌گیری در زمان (سیستم‌های ترتیبی)

- رباتیک و کنترل مسیر در طول زمان
- پیش‌بینی وضعیت بعدی در سیستم‌های دینامیکی

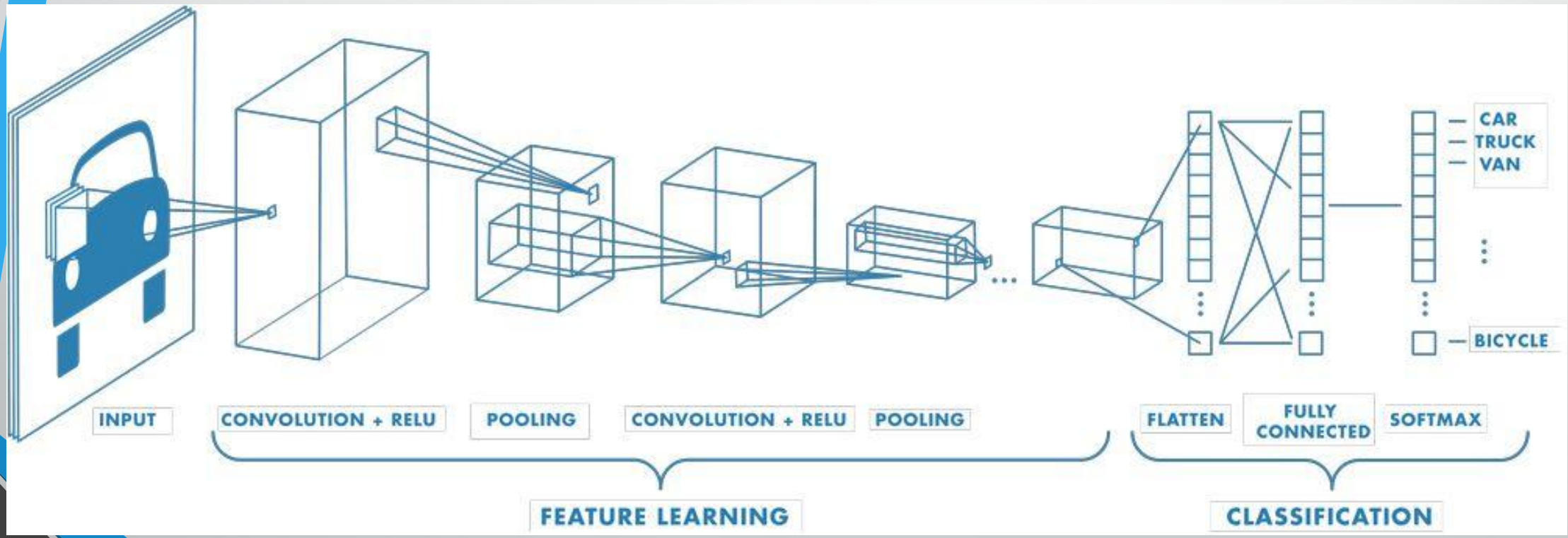
(گاه با Reinforcement Learning یا مدل‌های ترکیبی استفاده می‌شود.)

CNN

- ساختار CNN (Convolutional Neural Network)

CNNتصویر را با فیلترهای کوچیک تکه‌تکه اسکن می‌کند،

ویژگی‌ها را یاد می‌گیرد و در آخر بین کلاس‌ها انتخاب می‌کند.



شبکه عصبی کانولوشنی CNNچیست؟

CNNیک نوع شبکه عصبی مصنوعی است که مخصوص پردازش تصویر طراحی شده است.

مثلاً برای:

تشخیص گربه و سگ

تشخیص چهره

تشخیص تومور در تصاویر پزشکی

•چرا CNNساخته شد؟

•شبکه‌های عصبی معمولی ((ANNبرای تصویر مناسب نیستند

• چون:تصاویر داده‌های خیلی بزرگی هستند (مثلاً ۱۰۰۰×۱۰۰۰ پیکسل)

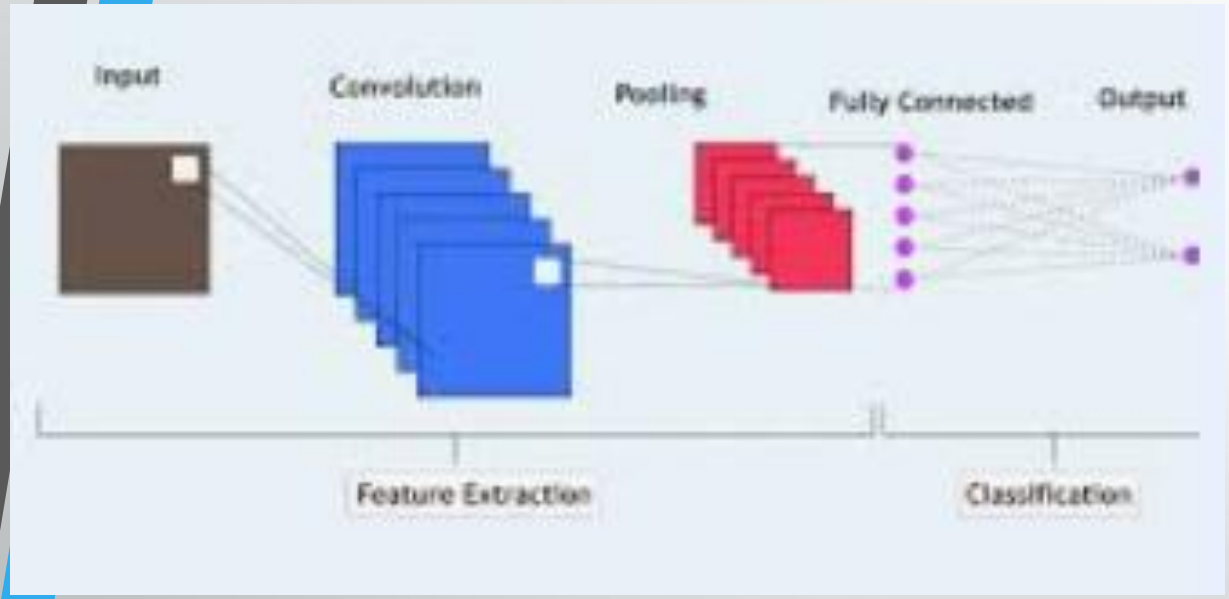
•تعداد پارامترها خیلی زیاد می‌شود

•محاسبات سنگین می‌شود

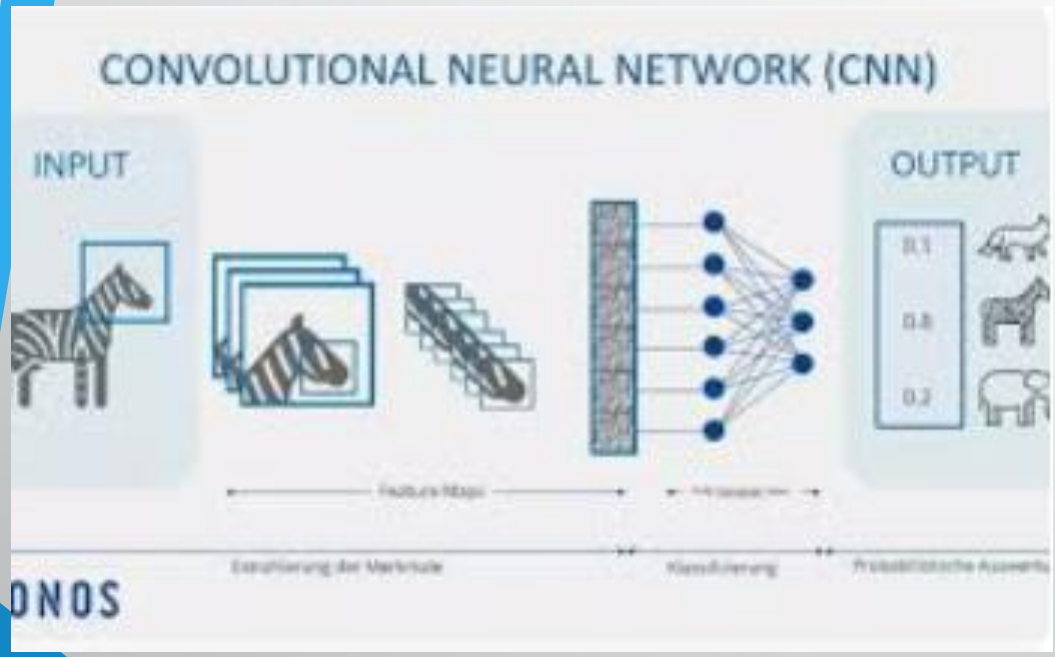
• CNN این مشکل را با تمرکز روی ویژگی‌های مهم تصویر حل می‌کند.

ساختار اصلی CNN

- یک CNN معمولاً از این لایه‌ها تشکیل شده:
  - Convolution Layer لایه کانولوشن
  - Activation معمولاً ReLU
  - Pooling Layer
  - Fully Connected Layer



• Convolution Layer لایه کانولوشن



در این مرحله:

- یک فیلتر کوچک (مثلاً ۳×۳) روی تصویر حرکت می‌کند
- هر جا الگوی خاصی ببیند آن را تشخیص می‌دهد

این فیلترها یاد می‌گیرند:

- لبه‌ها
- خطوط
- گوشه‌ها
- بافت‌ها

نتیجه: تولید یک تصویر جدید به نام Feature Map

Activation Function (ReLU)

بعد از کانولوشن، از تابع ReLU **استفاده می‌شود:**

$$ReLU(x) = \max(0, x)$$

یعنی:

- **اگر عدد مثبت باشد → همان را نگه می‌دارد**
- **اگر منفی باشد → صفر می‌شود**

باعث می‌شود مدل غیرخطی شود و بهتر یاد بگیرد.

فرض کنید خروجی یک فیلتر کانولوشن در یک قسمت از تصویر این اعداد باشه:

[۲, −۱, ۵, ۰, −۳]

وقتی تابع ReLU **رو روی این اعداد اعمال می‌کنیم:**

۲ مثبت است ← خروجی ۲

−۱ منفی است ← خروجی ۰

۵ مثبت است ← خروجی ۵

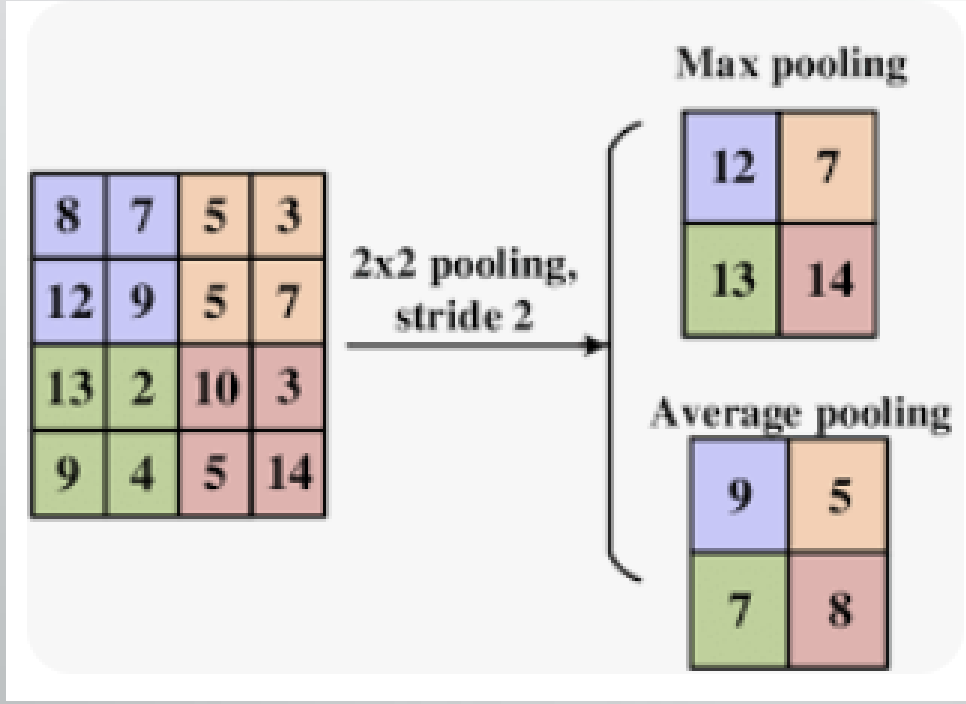
۰ صفر است ← خروجی ۰

−۳ منفی است ← خروجی ۰

پس، خروجی بعد از اعمال ReLU **میشه:**

[۲, ۰, ۵, ۰, ۰]

Pooling Layer (کوچک‌سازی تصویر)



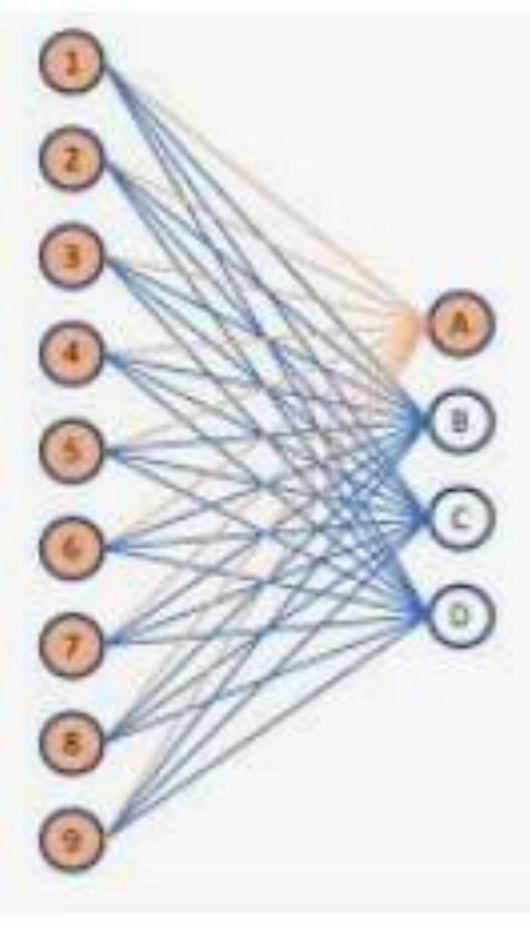
- اندازه تصویر را کوچک می‌کند
- اطلاعات مهم را نگه می‌دارد
- سرعت پردازش را بالا می‌برد

معمولاً از Max Pooling استفاده می‌شود:

از هر قسمت فقط بزرگ‌ترین عدد را نگه می‌دارد

مزیت: کاهش محاسبات و جلوگیری از Overfitting

Fully Connected Layer



در انتهای شبکه:

- **تمام ویژگی‌های استخراج شده جمع می‌شوند**
- **شبکه تصمیم نهایی می‌گیرد**

مثلاً:

۰/۹ → گربه

۰/۱ → سگ

کاربردهای اصلی CNN

۱) بینایی کامپیوتر (تصاویر)

- طبقه‌بندی تصویر
- تشخیص شیء
- بخش‌بندی تصویر

۲) تشخیص چهره و اثر انگشت

- شناسایی چهره
- تشخیص ویژگی‌های صورت
- در بعضی سیستم‌ها، تشخیص روی تصویر/دوربین

۳) تشخیص دست‌خط

- خواندن شماره‌ها و حروف از روی تصویرهای دست‌نویس
- OCRهای مبتنی بر CNN

۴) پزشکی و تصاویر پزشکی

- تشخیص بیماری از روی عکس (مثل MRI/CT/X-ray)
- شناسایی نواحی غیرعادی در تصاویر پزشکی

۵) تحلیل ویدئو (در ترکیب با سایر مدل‌ها)

خودِ CNN معمولاً برای تصویر است، ولی در سیستم‌های ویدئویی به عنوان جزء اصلی استفاده می‌شود:

- تشخیص اکشن/رویداد
- استخراج ویژگی از فریم‌ها (و بعد مدل زمانی مثل LSTM/Transformer رویشان کار می‌کند)

۶) کاربردهای صنعتی و امنیتی

- کنترل کیفیت در کارخانه با دید ماشین
- تشخیص نقص روی قطعات
- تشخیص فعالیت‌ها در دوربین‌های نظارتی