

মেশিন লার্নিং (১): কি এবং কেন শিখবো?

Machine learning Bengali tutorial



Sharif Hasan

• July 22, 2020

সর্বশেষ আপডেট April 2, 2021

7

1,971

পড়তে 3 মিনিট লাগতে পারে

অনেক দিন পরে আবার লেখতে বসলাম। মেশিন লার্নিং নিয়ে একটি কোর্স করছিলাম। এখন ভাবলাম ব্লগটাতে লেখি এটা নিয়ে। তাই লেখা শুরু করলাম। এই লেখাটা ইনশাল্লাহ্‌ চেইন আকারে লিখবো এবং সুপারভাইসড লার্নিং টা কভার করার চেষ্টা করবো।

মেশিন লার্নিং (Machine learning or ML) কি

আমরা প্রত্যেকদিন নানাভাবে না জেনেই মেশিন লার্নিং আপ্লিকেশন গুলো ব্যবহার করছি। গুগলে আমাদের প্রত্যেকটা সার্চ রেজাল্ট এর সঙ্গে মেশিন লার্নিং জড়িত। সার্চ ইঞ্জিন গুলো মেশিন লার্নিং ব্যবহার করে ওয়েব পেইজ গুলোকে র‍েঙ্ক করে। ফেইসবুক এর অটোমেটিক ফটো ট্যাগিং সিস্টেমও এর ব্যবহার করে। এরকম আরও অনেক ব্যবহার বলা যাবে খুব সহজেই।

মেশিন লার্নিং কি সে বিষয়ে আরথার সেমুয়েল বলেছেন



The field of study that gives computers the ability to learn without being explicitly programmed.





আর্থার সেমুয়েল

সুতরাং কম্পিউটার কে নির্দিষ্টভাবে প্রোগ্রাম না করেই কম্পিউটার কে কোন কিছু শিখানোর সক্ষমতা দেয়ার উপায়ই মেশিন লার্নিং।

টম মিশেল মেশিন লার্নিং এর একটি আধুনিক সংজ্ঞা দিয়েছেন। তিনি বলেছেন,



*A computer program is said to learn from experience **E** with respect to some class of tasks **T** performance measure **P**, if its performance at tasks in **T**, as measured by **P**, improves with experience **E**.*

কার্ড খেলার উদাহরণ এর মাধ্যমে বুঝিয়ে বলা যাক

1. **E** = কার্ড খেলার অভিজ্ঞতা (Experience)
2. **T** = কার্ড খেলার কাজ (Task)
3. **P** = পরবর্তী গেম এ কম্পিউটার জিতার সম্ভাবনা

সাধারণত মেশিন লার্নিং এর সমস্যা গুলোকে প্রধানত দুই ভাগে ভাগ করা হয়।

1. সুপারভাইসড লার্নিংঃ এই ধরনের সমস্যা সমাধানের জন্য আমাদের ডাটাসেট এ বলে দেয়া থাকে যে আমাদের টার্গেট আউটপুট কি হবে।
2. আনসুপারভাইসড লার্নিংঃ এই ধরনের সমস্যায় বলা থাকে না আউটপুট কি হবে। ডাটাসেট বিশ্লেষণ করে ডাটাসেট এর স্ট্রাকচার বের করে নিতে হয়।

Tom Mitchel

সুপারভাইসড মেশিন লার্নিং

সুপারভাইসড লার্নিং এর ক্ষেত্রে আমাদের একটি ডাটাসেট দেয়া থাকে এবং বলা থাকে আমাদের সঠিক আউটপুট কি হবে। এক্ষেত্রে অবশ্যই ইনপুট এবং অউটপুট এর সাথে রিলেশন থাকে। একটা ডাটাসেট এর উদাহরণ দেখি।

X_1	X_2	Y
100	200	800
300	400	1800
50	100	400
75	90	420

এখানে আমাদের একটি ডাটাসেট দেয়া আছে। প্রতিটা রোতে আমাদের X_1, X_2, Y যেখানে এদের মাঝে সম্পর্ক হলো $2 \times X_1 + 3 \times X_2 = Y$.

সুপারভাইসড মেশিন লার্নিং প্রবলেম গুলো দুই ভাগে বিভক্ত

1. রিগ্রেশন প্রবলেম
2. ক্লাসিফিকেশন প্রবলেম

রিগ্রেশন প্রবলেম গুলোতে আমরা রেজাল্ট গুলোকে একটা অবিচ্ছিন্ন অউটপুট হিসেবে প্রেডিক্ট করার চেষ্টা করি। যেমন উপরের ডাটাসেট এর দেখা যায় আমাদের আউটপুট Y X_1, X_2 এর একটি অবিচ্ছিন্ন ফাংশন।

অপরদিকে আমরা ক্লাসিফিকেশন প্রবলেম প্রবলেম করার সময় আমরা চলক গুলিকে একটি বিচ্ছিন্ন ক্যাটেগরিতে ম্যাপ করার চেষ্টা করি। যেমন, আজ বৃষ্টি হবে কি হবে না, কোন ছবিতে কার কার মুখ দেখা যায় ইত্যাদি।

উদাহরণ ১ঃ

ধরেন আমাদেরকে একটি হাউজিং কোম্পানির নির্দিষ্ট সাইজ এর হাউস এর জন্য দাম দেয়া আছে। আমাদেরকে কোন নতুন হাউজের অজানা সাইজের জন্য তার দাম প্রেডিক্ট করতে হবে। সুতরাং এটা একটা রিগ্রেশন প্রবলেম।

আমরা এই উদাহরণটিকে একটি ক্লাসিফিকেশন প্রবলেম এ রূপান্তর করে দিতে পারি। এজন্য আমাদের মেশিন লার্নিং মডেল কোন ধরনের দাম প্রেডিকশন করবে না। সে যা করবে তা হলো কাস্টমার যে দাম অফার করবে সে দাম এ বিক্রি করবো কি করবো না তা বলে দিবে।

উদাহরণ ২ঃ

1. রিগ্রেশন – একটি লোকের ছবি দেয়া আছে, আমাদের বলতে হবে লোকটির বয়স কতো।
2. ক্লাসিফিকেশন – আমাদের বলতে হবে উক্ত ছবিটি তে যে মানুষটি আছে সে ছেলে না মেয়ে।

নোটঃ এখন বুঝতে একটু সমস্যা হলেও পরে সমাধান হয়ে যাবে আশা করি।

আন-সুপারভাইসড লার্নিং (**Unsupervised learning**)

এইধরনের সমস্যার ক্ষেত্রে আমাদের অউটপুট কি হবে তা অল্প জানা বা পুরোপুরি অজানা থাকে। এক্ষেত্রে আমাদের ডাটাসেট এর মধ্যে যেসব ডাটা আছে তাদের মধ্যে

একটি স্ট্রাকচার খুঁজে বের করতে হয় যেখানে ডাটাসেট কোন চলকের কি প্রভাব তা নিয়ে প্রায় কোনও ধারণা থাকে না। আমরা ডাটাসেট এর চলক গুলোর পারস্পরিক সম্পর্কের উপর ভিত্তি করে আমরা ডাটা এর গঠন খুঁজে পেতে পারি আনসুপারভাইসড লার্নিং অ্যালগরিদম ব্যবহার করে।

উদাহরণঃ ১,০০০,০০০ টি বিভিন্ন ধরনের জিন এর ডাটাসেট দেয়া আছে, আমাদের এই জিন গুলোকে স্বয়ংক্রিয় ভাবে গ্রুপ করতে হবে যাতে একই গ্রুপের জিন গুলো কোন ভাবে একে অপরের সাথে সম্পর্কিত (যেমনঃ জীবনসীমা, অঞ্চল ইত্যাদি) হয়।

মেশিন লার্নিং এর জন্য দরকারি কিছু সংজ্ঞাঃ

1. মডেল: মডেল হলো আমরা কোন সমস্যা যেভাবে সমাধান করি তার একটি গাণিতিক উপস্থাপন। মেশিন লার্নিং মডেল তৈরি করার জন্য আমাদেরকে একটি ডাটাসেট দেয়া থাকবে যার থেকে মেশিন লার্নিং অ্যালগরিদম গুলো শিখবে।
2. ML অ্যালগরিদম: মেশিন লার্নিং অ্যালগরিদম হলো গণিত এবং যুক্তি এর সমষ্টি। প্রদত্ত ডাটার উপরে ভিত্তি করে মেশিন লার্নিং অ্যালগরিদম গুলো পারফর্ম করে। যেমনঃ লিনিয়ার রিগ্রেশন, লজিস্টিক রিগ্রেশন, নিউরাল নেটওয়ার্ক ইত্যাদি।
3. ট্রেনিং: ট্রেনিং ML অ্যালগরিদম গুলোকে শিখানোর প্রসেস। ট্রাইনিং এর সময় মেশিন লার্নিং অ্যালগরিদম গুলো ডাটা এবং আউটপুট এর মধ্যে প্যাটার্ন খুঁজে বের করে।

কৃত্রিম বুদ্ধিমত্তা এবং মেশিন লার্নিং: পার্থক্য কি?

কৃত্রিম বুদ্ধিমত্তা অপেক্ষাকৃত ভাবে একটি বড় ধারণা যার মাধ্যমে মানব চিন্তা ধারা মেশিনে সিমুলেট করার চেষ্টা করা হয় সেখানে Machine learning কৃত্রিম বুদ্ধিমত্তার

একটি সাবসেট যার মাধ্যমে ডেটা থেকে মেশিনকে শিখানো হয়।

আজকে এই পর্যন্তই। আরও কিছু সংজ্ঞা আছে, সেগুলো আস্তে আস্তে লিখবো। আগামী লেখায় আমরা লিনিয়ার রিগ্রেশন এর কস্ট ফাংশন নিয়ে আলোচনা করবো। কারও কোন জায়গায় কনফিউশন থাকলে নিশ্চিত্তে কमेंট করুন। আপাতত বিদায়।

আরও পড়ুনঃ [গ্রাফঃ বিএফএস \(BFS\) গ্রাফ ট্রাভার্সাল অ্যালগরিদম](#), [Machine learning – wikipedia](#)

লেখাটি কেমন লেগেছে আপনার?

রেটিং দিতে হার্টের উপর ক্লিক করুন।



গড় রেটিং 3.8 / 5. মোট ভোট: 18

#AI

#Machine learning

#মেশিন লার্নিং

7 টি মন্তব্য



Sanzid

July 22, 2020 at 11:52 PM

ভাল ছিল। আশাবাদী আরও Deep কিছু পাব পরবর্তীতে।

Reply

Pingback: মেশিন লার্নিং (২): লিনিয়ার রিগ্রেশন, মডেল রিপ্রেজেন্টেশন ও কস্ট ফাংশন

Pingback: মেশিন লার্নিং (২): লিনিয়ার রিগ্রেশন, মডেল রিপ্রেজেন্টেশন ও কস্ট (Cost) ফাংশন

Pingback: মেশিন লার্নিং (২): লিনিয়ার রিগ্রেশন, মডেল রিপ্রেজেন্টেশন ও কস্ট (Cost) ফাংশন

Pingback: কস্ট ফাংশনের অন্তরীকরণ এবং গ্র্যাডিয়েন্ট ডিসেন্ট - শরিফ হাসানের ব্লগ

Pingback: মেশিন লার্নিং (3): কস্ট ফাংশনের অন্তরীকরণ এবং গ্র্যাডিয়েন্ট ডিসেন্ট - শরিফ হাসানের ব্লগ



Odri Chakma

September 5, 2021 at 7:42 AM

Khub Bhalo LAAGLO..

Reply