

دانش

صفر و یک

فصلنامه دانشجویی علوم کامپیوتر



} کاربردهای زبان‌های برنامه‌نویسی مختلف

الگوریتم A*

جایگاه IoT در زندگی هوشمند

{ حتی کد بد هم می‌تواند کار کند.

شماره ۱
بهار ۱۴۰۱



0001

اسم لغت



شناسنامه

صفر و یک، فصلنامه دانشجویی علوم کامپیوتر
دانشگاه الزهرا (س) - تهران
شماره اول - بهار ۱۴۰۱
صاحب امتیاز: انجمن علمی علوم کامپیوتر دانشگاه الزهرا
مدیر مسئول: زهرا رفیع
سرمدیر: زهرا رفیع، سادنه صادقی اسکوتی
استاد راهنما: دکتر بیتا شمس



تعمین محتوا و تدوین

زهرا رفیع، سادنه صادقی اسکوتی، کیمیا جنکی میاردان، مانا سامانی، اعظم سادات کامرواف، مبینا عربزاده، سحر حقیقت، مبینا محمدحسینی لالائی، اسما بنکدارزاده، مبینا سادات شریعتپناهی، سما مرادی، ریحانه محمدزاده، سوگند روستایی، مبینا خیری، فاطمه آریانی کریمی، زهرا صدوق، کیمیا بارسم، شقایق روزمه



ویراستاری

محدثه علیخانی، مریم خلیلی، مبینا عربزاده، سمانه ملاکریمی، سادنه صادقی اسکوتی، مهدیه خرمند حاجی‌باشی، زهرا رفیع



طراحی گرافیک

مهدیه اسمعیلزاده، عارفه اخگری، اسما بنکدارزاده، ریحانه ضیفمی، نیکا رضایی



طراحی جلد و صفحه‌آرایی

شادک بیانی‌پور



پاهاهای ارتباطی

✉ alzcssjournal01@gmail.com

📍 [@AlzCSSJournal01](https://www.instagram.com/AlzCSSJournal01)

🌐 [AlzCSSJournal.01](https://www.facebook.com/AlzCSSJournal.01)

۶ ماجراهای من و کامپیوترم

یک نود و نهی در الزهرا گزارش می‌دهد.
بخش ۱: چس شد که اینطور شد؟

۷ علوم کامپیوتر

- ۹ علوم کامپیوتر، پلی بین دانش ریاضی و کامپیوتر
- ۱۱ تفاوت رشته علوم کامپیوتر با رشته‌های مشابه
- ۱۲ ۷ تخصص برتر علوم کامپیوتر
- ۱۵ یک مسئله حل‌نشده در علوم کامپیوتر
- ۱۶ از چرتکه تا کامپیوتر
- ۱۸ الگوریتم فشرده‌سازی کدگذاری هافمن
- ۲۰ بررسی کاربردهای زبانهای برنامه‌نویسی مختلف
- ۲۲ مهارت‌های مورد نیاز در علوم کامپیوتر
- ۲۴ متاورس، نسل بعدی اینترنت
- ۲۶ معرفی کتاب و دوره در علوم کامپیوتر
- ۲۷ انتخاب رشته مناسب علوم کامپیوتر برای کارشناسی ارشد

۳۰ داده‌کاوی

- ۲۹ داده‌کاوی، کشف دانش در پایگاه‌داده‌ها
- ۳۱ کاربردهای داده‌کاوی
- ۳۲ مزایا و محدودیت‌ها در داده‌کاوی
- ۳۴ کاربردهای ابزارهای تحلیلی داده‌کاوی و متن‌کاوی در چاپکی سازمان‌های مراقبت بهداشتی و درمانی
- ۳۶ الگوریتم k-means
- ۳۸ زنان در علم داده، مارچ ۲۰۲۲
- ۳۹ دروس و ابزارها در داده‌کاوی
- ۴۱ مراحل کاوش داده و مهارت‌های ضروری
- ۴۲ معرفی کتاب و دوره در داده‌کاوی

۴۲

هوش مصنوعی

۴۴ هوش مصنوعی =! هوش طبیعی

۴۵ زیرشاخه‌ها و کاربردهای هوش مصنوعی

۴۸ چالش‌های حل‌شده و حل‌نشده در هوش مصنوعی

۴۹ هوش مصنوعی و ژنتیک

۵۱ الگوریتم A*

۵۲ دروس و ابزارها در هوش مصنوعی

۵۴ مسیر یادگیری هوش مصنوعی و مباحث مقدماتی

۵۵ معرفی کتاب و دوره در هوش مصنوعی

۵۶

بیوانفورماتیک

۵۷ بیوانفورماتیک یا زیست‌داده‌ورزی!

۵۹ کاربردها و زمینه‌های بیوانفورماتیک

۶۰ حل یک مثال از الگوریتم‌های ترکیبیاتی

۶۲ بررسی بیوانفورماتیکی توالی ژن ایتکزاز ویروس‌های HIV-1 در بیماران ایرانی

۶۳ دروس و ابزارها در بیوانفورماتیک

۶۶ معرفی کتاب و دوره در بیوانفورماتیک

۶۷

اینترنت اشیا

۶۸ اینترنت اشیا، هویت‌بخشی دیجیتال به اشیاء بی‌جان

۷۱ جایگاه IoT در زندگی هوشمند

۷۲ چالش‌های حل‌شده و حل‌نشده در اینترنت اشیا

۷۴ کاربرد اینترنت اشیا در کنترل بیماری کووید ۱۹

۷۶ الگوریتم RSA

۷۷ مهارت‌های مورد نیاز یک مهندس اینترنت اشیا

۷۸ دروس و ابزارها در اینترنت اشیا

۷۹ نقشه یادگیری مباحث اینترنت اشیا

۸۰ معرفی کتاب و دوره در اینترنت اشیا

۸۱

کد تمیز

حتی کد بد هم می‌تواند کار کند.

-خلاصه نکات چهار فصل نخست کتاب کد تمیز-

سخن سردییر

به نام خدا

همراهان صفر و یک سلام،

بسیار خرسندیم که با توکل بر خدا و همکاری و همدلی دوستان، با آغاز بهارِ قرن جدید، شروع تازه‌ای نیز برای فصل‌نامه صفر و یک رقم خورد.

در ابتدا لازم می‌دانم از تمام عزیزانی که با تمام توان در این فرصت اندک برای به ثمر رسیدن این نشریه تلاش کردند، کمال تشکر را داشته باشم.

از سرکار خانم دکتر بیتا شمس، استاد راهنمای نشریه، و سرکار خانم صدی، مسئول نشریات، که ما را در این مسیر راهنمایی کردند نیز بسیار سپاسگزاریم.

هدف از ایجاد این نشریه آشنایی بیشتر دانشجویان رشته علوم کامپیوتر و سایر همراهان عزیز با ماهیت و مسائل روز علم گسترده کامپیوتر می‌باشد.

در اولین شماره از صفر و یک سعی کردیم به طور خاص و متفاوت مطالبی که در فضای تیمی خود درباره علوم کامپیوتر و زیرشاخه‌های آن آموخته بودیم، از دیدگاه علمی بررسی و در قالب پنج بخش علوم کامپیوتر، داده‌کاوی، هوش مصنوعی، بیوانفورماتیک و اینترنت اشیا در اختیار شما عزیزان قرار دهیم.

امید است قدمی در راستای انتقال درست و اصولی مطالب و ارتقا سطح علمی شما عزیزان نسبت به رشته خودمان برداشته باشیم.

در آخر امیدواریم در آینده با تلاش و همکاری سایر دانشجویان شاهد ادامه یافتن و انتشار شماره‌های بیشتری از فصل‌نامه صفر و یک باشیم.

ماجراهای من و کامپیوترم یک نود و نه در الزهرا گزارش می‌دهد.

- بخش ۱: چه شد که این طوری شد؟ -



تیر ماه ۹۸ بود، من سال بعدش همون موقعها کنکور داشتم و باباحاجی داشت نفس‌های آخرش رو می‌کشید...

دفعه آخر که دیدمش بی‌جون بود و لب‌هاش می‌لرزیدن، دست‌هاش هم همین‌طور.

رفتم کنار تختش. وقتی منو دید بی‌رمق و نصف و نیمه گفت: دکتر، حال خوب، آدم‌ها...

بعدش، همین‌طور تا یکی دو روز خواب و بیدار و مریض و بعدشم که، تمام!

یک دنیایی رو با خودش برد زیر خاک...

فکرش رو هم نمی‌کردم که یک روزی، نقش اول همه فیلم‌های محبوب زندگیم مثل یک ستاره خاموش بشه.

که سینمای همون سرزمین تم که قبل‌تر ذکرش رفت، دیگه مادام‌العمر تعطیل بشه.

ولی شد!

دیگه خالی خالی بودم، یک خلا بکر. با این حساب ولی چون با باباحاجی قول و قرار گذاشته بودیم، درس خوندم که دکتر بشم...

چند ماهی گذشت. اواخر بهمن بود و اوایل اسفند که کرونا اومد. یک جور طاعونی که افتاد به جون زندگی آدما ولی در قرن بیستم!

مرداد ۹۹ بود که کنکور با تمام تمایز از اون تعریف فیکسش برگزار شد.

خیلی‌ها که دیگه از همون اسفند به خاطر جناب کرونا درس و کنکور رو ول کردن، چون یا درگیر کرونا شدن یا درگیر طفیلی اون، یعنی قرنطینه و مشکلات ناشی ازش، که ملموس‌ترینش افسردگی بود.

عده‌ای هم از شرکت تو کنکور که اجتماع بزرگی بود ترسیدن، دوباره

از وقتی که یادم میاد درس خوندم...

با خودم تو عالم بچگی فکر می‌کردم - حالا فکر نه به مثابه اون سمینار رسمی عقل و منطق، سوا از معنی فلسفی‌بی که می‌تونه داشته باشه - درس خوندم می‌تونه اون نردبون ترقی معروف باشه.

وقتی ابتدایی بودم، درس خوندم که برم به راهنمایی خوب...

وقتی راهنمایی بودم، درس خوندم که برم به دبیرستان خوب...

و وقتی دبیرستان بودم، درس خوندم که بتونم کنکور قبول بشم و برم به دانشگاه خوب..

همین‌طور هی خوندم و خوندم و این خوندم با من بزرگ شد و شد اصل پایداری زندگی من.

این وسط پدر بزرگ همیشه تشویق می‌کرد، نه این‌که بقیه تشویق نکنان، نه!

اما باباحاجی - پدر بزرگ رو می‌گم - به طور بهترترتری تشویق می‌کرد.

دلم می‌خواست هر سال باباحاجی اولین کسی باشه که کارنامه‌مو بهش نشون می‌دم.

همیشه چشمم به دهنش بود تا بهم کارت صدآفرین بده یا نمره انضباط - اونم فقط بیست! -

حالا تشویقش چی بود؟

تشویقش یه لبخند سحرآمیز بود - برای من کارزماتیک‌تر از لبخند مونالیزا - و این جمله طلایش که تو سرزمین تم، سر در شهر نوشته بودنش و همه سربازان وظیفه، صبح، ناشتا، قبل از هر کاری باید بهش احترام نظامی می‌داشتن:

تو می‌تونی دکتر بشی، مطمئنم! درس بخون بابا، دکتر بشی حال آدما رو خوب کنی...

بمخاطر همین جناب کرونا.

من ولی کنکور دادم و تونسته بودم که دکتر بشم...

ولی نه اون دکتری که شما الان متصور می‌شید. از اولشم باباحاجی منظورش از دکتر شدن به منظور استعاری بود، یعنی به اصلی که کادوپیشش کرده باشن تا خودت بازش کنی و کشفش کنی...

آخه خب مگه فقط دکترا حال آدما رو خوب می‌کنن؟ می‌دونید منظورم چیه؟ مثلاً فرض کنیم شما الان از صبح دانشگاه بودید و پشت سر هم تا ساعت هفت شب کلاس داشتید، خسته‌اید، گرسنه‌اید و داغون بعد از به روز سخت رسیدید خونه.

مادرتون در خونه رو براتون باز می‌کنه و وقتی می‌بینتون گل از گلش می‌شکفه و شروع می‌کنه قریون صدقه قد و بالای رعنائتون رفتن، بعدشم به

لیوان چای زعفرونی لبریز و لبسوز و لبدوز می‌ذاره جلوتون.

حتی فکر کردن بهش حالتونو خوب می‌کنه، نه؟

پس این‌جا با استدلال من، مادر، همون دکتری می‌شه که راجع بهش حرف زدیم. دقیقاً همون دکتری که مد نظر باباحاجی‌ام بود. سرتونو درد نیارم...

من دانشگاه الزهرا، کامپیوتر قبول شدم و با خودم قرار گذاشتم انقدر تلاش کنم تا بتونم حال خیلی‌ها رو خوب کنم با رشته‌ای که قبول شدم و حالا دیگه قراره بشه تخصصم...

قرار گذاشتم بشم بهترین مهندس دکتر!

بله! من و کامپیوترم دیگه ماجراها خواهیم داشت.

علوم کمپیوٹر



0001



علوم کامپیوتر، پله بین دانش ریاضه و کامپیوتر

علم کامپیوتر عبارت است از تئوری و عملی که به شما امکان می‌دهد کامپیوتر را طوری برنامه‌ریزی کنید تا آنچه را که می‌خواهید انجام دهد، ابزاری که به شما کمک می‌کند داستانی را تعریف کنید یا اتفاقی را با فناوری بسازید. رشته‌ای که بر پایداری در حل مسئله تأکید دارد، مهارتی که در همه رشته‌ها قابل اجرا است و باعث رشد شغل و نوآوری در تمام بخش‌های نیروی کار می‌شود. مهارتی که به دانش آموزان می‌آموزد که چگونه از رایانه برای خلق کردن استفاده کنند نه صرفاً مصرف.

و این درحالی است که افراد علوم کامپیوتر را با موارد زیر اشتباه می‌گیرند: یادگیری نحوه تایپ یا استفاده از ماوس، یادگیری استفاده از نرم افزار پردازش کلمه، صفحه گسترده یا ارائه (مانند Word، PowerPoint، Google Docs و Drive)، آموزش ساخت یا تعمیر کامپیوتر، بازی‌های ویدیویی، مهارت‌های متوازن هوشمند (SBAC).

اما ما سعی داریم در این نشریه به بررسی دانشگاهی تعاریف و رشته‌مان بپردازیم، و با خودداری از گسترش باورهای غلط سعی در توسعه موارد صحیح در رشته علوم کامپیوتر را داریم.

یادگیری علوم کامپیوتر با سرعت باورنکردنی وارد نیازها و زندگی روزمره انسان‌ها شد.

اولین دانشگاهی که مدرک خود را در رشته علوم کامپیوتر معرفی کرد، دانشگاه «کمبریج» در سال ۱۹۵۳ بود. اولین گروه علوم کامپیوتر در ایالات متحده در دانشگاه «پر دو» در سال ۱۹۶۲ تشکیل شد.

همچنین رشته علوم کامپیوتر برای اولین بار در ایران، در سال ۱۳۵۸ به عنوان یک رشته مادر وارد دانشگاه تهران شد. اما در سال ۱۳۶۵ با راه اندازی دو شاخه اصلی مهندسی کامپیوتر، این رشته از لیست رشته‌های انتخابی دانشگاه حذف شد. اما خوشبختانه وزارت علوم با بیان اینکه نقش و اهمیت این رشته بسیار بالاست، در سال ۱۳۷۷ بار دیگر این رشته را، البته این بار وارد اکثر دانشگاه‌های ایران کرد.

رشته علوم کامپیوتر برای مطالعه نحوه محاسبات به وجود آمده و موثرترین راه ارتباطی بین ریاضیات و کامپیوتر است. مهم‌ترین هدف این رشته دستیابی به بهترین روش برای حل یک مسئله در کمترین زمان و با بالاترین دقت است. دغدغه اصلی علوم کامپیوتر این است که «چه چیزی می‌تواند» و «چه چیزی نمی‌تواند» خودکار شود. علوم کامپیوتر در علوم دیگر نیز تاثیر گسترده‌ای داشته است. برای مثال شبیه‌سازی پدیده‌های فیزیکی و فیزیک محاسباتی، استفاده وسیع از ربات‌ها در تحقیقات فضایی و هر محیطی که امکان حضور انسان در آن مکان‌ها نیست، تشخیص بیماری‌ها به کمک الگوریتم‌های کامپیوتری و ... نمونه‌هایی از این تاثیر می‌باشند.

پس با توجه به توضیحات بالا، علوم کامپیوتر مطالعه کامپیوترها و سیستم‌های محاسباتی است. برخلاف مهندسان برق و کامپیوتر، دانشمندان کامپیوتر بیشتر با نرم‌افزارها و سیستم‌های نرم‌افزاری سروکار دارند. این شامل تئوری، طراحی، توسعه و کاربرد آنها می‌شود.

زمینه‌های اصلی تحصیل در علوم کامپیوتر شامل «هوش مصنوعی، سیستم‌ها و شبکه‌های کامپیوتری، امنیت، سیستم‌های پایگاه داده، تعامل کامپیوتری انسانی، بینایی و گرافیک، تحلیل عددی، زبان‌های برنامه‌نویسی، مهندسی نرم‌افزار، بیوانفورماتیک و تئوری محاسبات» است. همچنین این رشته شامل مطالعه الگوریتم‌ها و ساختارهای داده، طراحی کامپیوتر و شبکه، مدل‌سازی داده‌ها و فرآیندهای اطلاعاتی و هوش مصنوعی می‌باشد. مهم‌ترین هدف این رشته پیدا کردن بهترین الگوریتم‌ها برای حل مسائل در کم‌ترین زمان و با کم‌ترین خطا می‌باشد.

نخستین زمره‌های پیدایش رشته علوم کامپیوتر از سال ۱۹۲۸ شروع شد. بهتر است بگوییم که این تاریخ، مرحله و نقطه‌ای نورانی برای تولد این رشته به حساب می‌آید. آن روزها تعریفی از رشته علوم کامپیوتر وجود نداشت. درست در سال ۱۹۲۸ بود که ریاضیدان آلمانی، «دیوید هیلبرت»، ۲۳ مسئله طلایی و حل نشده در ریاضیات را با عنوان «مسائل اصلی علم ریاضیات در آینده» مطرح کرد. او این سوالات اساسی را با هدف حل شدن توسط ریاضی‌دانان مشهور آن زمان به اشتراک گذاشت. سه مورد از این مسائل تاثیر گسترده‌ای بر شکل‌گیری علوم کامپیوتر و تاسیس این شاخه از دانش بشری داشت.

در سال ۱۹۳۱ منطق‌دانی به نام «کرت گدل» توانست به دو پرسش از این سه پرسش مطرح شده، پاسخ بدهد. درست ۵ سال بعد، در سال ۱۹۳۶ فردی به نام «آلن تورینگ» ریاضیدان، دانشمند رایانه، منطق‌دان، فیلسوف و زیست-ریاضیدان انگلیسی، که بعدها به پدر علوم کامپیوتر شهرت یافت، با معرفی یک مدل ریاضی از یک ماشین محاسبه‌گر (ماشین تورینگ) آخرین پرسش باقی‌مانده از این پازل ۳ قسمتی را حل کرد.

بسیاری، این اتفاقات را انقلابی در «نظریه محاسبه» دانسته و معتقد هستند نخستین گام‌های بزرگ در راستای شکل‌گیری و پیشرفت علوم کامپیوتر با این رخدادها برداشته شده است.

از طرفی وقوع جنگ جهانی دوم در پیشرفت علوم کامپیوتر نقش موثری داشت. چراکه در آن زمان ترانزیستور اختراع شده و این دستگاه نیاز شدید به ابزار محاسباتی خودکار داشت. در همان زمان تصمیم گرفته شد تا برای ساخت ماشین‌های محاسبه‌گر، سرمایه‌گذاری گسترده‌ای انجام شود. در واقع اگر پیشرفت‌های دهه ۱۹۵۰ در زمینه ساخت کامپیوتر نبود، هیچ وقت علوم کامپیوتر به عنوان یک شاخه مستقل در میان رشته‌ها و علوم شکل نمی‌گرفت.

کنار هم قرارگرفتن ماشین‌های محاسبه‌گر از یک سو و پیشرفت علم الکترونیک از سوی دیگر، باعث شد دانشمندان بتوانند ماشین‌های الکترونیکی، که معادل مدل‌های ریاضی موجود بود بسازند که می‌توانستند محاسبات ریاضی را با دقت و سرعت بالایی انجام دهند. اما علوم کامپیوتر هنوز نیاز به فرصت داشت تا خود را بهتر نمایان کند.

در آن زمان نیازهایی در جهان بود که در اثر گذر زمان رو به افزایش بود و از سویی دیگر کامپیوتر، این ابزار محاسبه‌گر قوی، روز به روز توانایی بیشتری برای پاسخگویی به نیازهای بشری می‌یافت و در نتیجه نیاز به

تفاوت رشته علوم کامپیوتر با رشته‌های مشابه

تفاوت رشته علوم کامپیوتر با دیگر رشته‌ها:

علوم کامپیوتر به عنوان بخشی از یک خانواده از پنج رشته جداگانه و در عین حال مرتبط به هم در نظر گرفته می‌شود: مهندسی کامپیوتر، علوم کامپیوتر، سیستم‌های اطلاعات، فناوری اطلاعات و مهندسی نرم‌افزار. اگرچه دانستن نحوه برنامه نویسی برای مطالعه علوم کامپیوتر ضروری است، اما تنها یکی از عناصر این رشته برنامه نویسی است. دانشمندان کامپیوتر الگوریتم‌هایی را برای حل برنامه‌ها و مطالعه عملکرد سخت افزار و نرم‌افزار کامپیوتر، طراحی و تجزیه و تحلیل می‌کنند. مشکلاتی که دانشمندان کامپیوتر با آن مواجه می‌شوند عبارت‌اند از: مسائل انتزاعی، تعیین اینکه چه مسائلی را می‌توان با رایانه‌ها حل کرد، پیچیدگی الگوریتم‌هایی که آنها را حل می‌کند، طراحی برنامه‌هایی که عملکرد خوبی بر روی دستگاه‌های دستی دارند، آسان بودن برای استفاده‌ی برنامه‌ها و اینکه تدابیر امنیتی را به درستی رعایت کنند.

الف) تفاوت رشته علوم کامپیوتر با مهندسی کامپیوتر: برای مقایسه خیلی کوتاه دو رشته مهندسی کامپیوتر و علوم کامپیوتر می‌توان گفت که در مهندسی سعی می‌شود یک ارتباط منطقی بین برق و کامپیوتر ایجاد شود ولی در رشته ی علوم کامپیوتر سعی در ایجاد یک ارتباط منطقی بین ریاضی و کامپیوتر می‌شود. پس اگر علاقمند به تئوری محاسبات و تحلیل و تجزیه داده‌ها دارید رشته علوم کامپیوتر می‌تواند انتخابی مناسب برای شما باشد. در اصل این رشته‌ها اغلب با هم هم‌پوشانی دارند، اما تفاوت‌های بین آنها قابل توجه است. اگر بتوانیم آنها را در یک جمله ساده خلاصه کنیم، می‌گوییم که علوم کامپیوتر با داده‌ها و الگوریتم‌ها سروکار دارد در حالی که مهندسی کامپیوتر بر روی سخت‌افزار و سیستم‌عامل متمرکز است. علوم کامپیوتر به عنوان شاخه‌ای از ریاضیات شروع شد، در حالی که مهندسی کامپیوتر به عنوان شاخه‌ای از مهندسی برق آغاز شد. و هنوز هم توسط برخی در نظر گرفته می‌شود.

**Computer
Science**

VS

**Computer
Engineering**

همچنین اگر نیازمند ساخت ماشین‌های فیزیکی قادر به برقراری ارتباط با انسان نبودیم، احتمالاً علوم کامپیوتر یک شاخه مخفیانه از ریاضی باقی می‌ماند. با توجه به اینکه بیشتر تحقیقات دانشگاهی، کم‌وبیش فرآیندی است، برای غربالگری اطلاعات انبار کاه برای یافتن سوزنی از اطلاعات مفید، پیشرفت‌های علوم کامپیوتر احتمالاً به ارائه ابزارهایی به محققان که در پردازش کارآمدتر مسائل و داده‌ها کمک می‌کند؛ منجر می‌شود - ارائه الگوریتم‌های بهتر، رایانه‌های بهتر و سریع‌تر و...-.

بدیهی است که هر چه انبار کاه در یک رشته خاص بزرگ‌تر باشد، علوم کامپیوتر به پیشرفت‌های بیشتری کمک خواهد کرد.

بله، پیشرفت در علوم کامپیوتر بر سایر رشته‌های دانشگاهی تاثیر می‌گذارد. ولی تقریباً تمام زمینه‌ها صرفاً به پیشرفت در علوم کامپیوتر وابسته نیستند، اما تحت تاثیر پیشرفت این رشته قرار می‌گیرند. اگر به صورت خلاصه و فقط با بیان عنوان‌ها این مطلب را بیان کنیم؛ می‌توان گفت که علوم کامپیوتر به طور کلی بدین‌گونه روی تمامی رشته‌ها تاثیر خواهد داشت و راه را آسان‌تر خواهد کرد:

- به طور مستقیم نیازها را بر آورده می‌کند.
- توانمند سازی افرادی که اغلب نادیده گرفته می‌شوند.
- هموار کردن راه برای جهانی عادلانه‌تر.
- تسریع مراقبت‌های بهداشتی.
- ادامه تحصیل.
- گسترش ارتباطات.
- پیش‌بینی و اجتناب از فجایع و رخدادها.
- تاثیر مثبت در هر بخش جامعه.

دانشجویان علوم کامپیوتر در مورد تئوری محاسبات، برنامه‌نویسی، محیط‌ها، الگوریتم‌ها، امنیت و انتقال داده‌ها و ... یاد می‌گیرند؛ درحالی‌که دانشجویان مهندسی کامپیوتر چگونگی توسعه، نمونه‌سازی و آزمایش ریزتراشه‌ها، مدارها، پردازنده‌ها، هدایت‌کننده‌ها و هر جزء دیگر مورد استفاده در دستگاه‌ها یا سیستم‌های هوشمند، لپ‌تاپ‌ها، سرویس‌دهنده‌ها، ابزارهای اینترنت اشیا و ... یاد می‌گیرند.

دانشمندان کامپیوتر با تئوری‌های محاسباتی، الگوریتم‌های قدرتمند و مدل‌های ریاضی کار می‌کنند که برای توسعه برنامه‌ها و سیستم‌های نرم‌افزاری استفاده می‌شود. از سوی دیگر مهندسان کامپیوتر، سخت افزار و سخت افزاری را که نرم‌افزارها و سیستم‌ها روی آن اجرا می‌شوند، توسعه می‌دهند.

ب) علوم کامپیوتر و مهندسی نرم‌افزار: تفکر عامه مردم این است که علوم کامپیوتر تفاوت چندانی با مهندسی نرم‌افزار ندارد اما این دیدگاه غلط است زیرا مهندسی نرم‌افزار به طور کامل مستقل از علم کامپیوتر است ولی نیاز کاری مهندسی نرم‌افزار آشنایی با علوم کامپیوتر است. اما به طور کلی علم کامپیوتر بیشتر با تئوری و روش‌هایی سروکار دارد که به کامپیوترها و سیستم‌های نرم‌افزاری مربوط می‌شود ولی مهندس نرم‌افزار با مسائل عملی مثل پیادسازی تحلیل و تولید نرم‌افزار سروکار دارند. مهندسی نرم‌افزار و علوم کامپیوتر رابطه خیلی نزدیکی باهم دارند و پیشرفت در علم کامپیوتر نیاز دارد که علم نرم‌افزار هم پیشرفت کند و لازم است که خیلی از کسانی که مهندسی نرم‌افزار را مطالعه می‌کنند از علوم کامپیوتر مطلع باشند.

علوم کامپیوتر در پایه و اساس خود رشته‌ای در مورد اطلاعات و پردازش اطلاعات است. اگر محاسبات کارهای فوق‌العاده مفیدی را انجام ندهند و

7 تخصص برتر علوم کامپیوتر

علوم کامپیوتر شامل طیف گسترده‌ای از مشاغل، عناوین شغلی و مسئولیت‌ها می‌شود. فرقی نمی‌کند شما به علوم، فناوری، مهندسی یا ریاضیات تعلق داشته باشید، علوم کامپیوتر می‌تواند شما را به حرفه‌ای که علاقه دارید متصل کند.

۷ تخصص برتر علوم کامپیوتر عبارتند از «هوش مصنوعی (AI)، یادگیری ماشینی (ML)، علم داده، تعامل انسان با کامپیوتر (HCI)، امنیت سایبری، پردازش ابری، رسانه‌های دیجیتال/تعاملی» که در ادامه به بررسی جزئیات بیشتری از آن‌ها می‌پردازیم.

۱. هوش مصنوعی (AI):

هوش مصنوعی امروزه بیش‌ترین تخصص را در علوم رایانه دارد، زیرا رایانه‌ها را قادر می‌سازد تا وظایف انسانی مانند ادراک بصری، تشخیص صدا، تصمیم‌گیری شناختی، تشخیص الگو و غیره را انجام دهند. از وسایل نقلیه خودران تا اتوماسیون خانگی و صنعتی؛ از سیستم‌های کنترل هوشمند گرفته تا تجهیزات پزشکی پیشرفته؛ از توسعه پیشرفته بازی تا پردازش زبان پیچیده، دامنه آن به قدری گسترده است که فرصت‌های شغلی و یادگیری برای همه و همه وجود دارد.

۲. یادگیری ماشینی (ML):

یادگیری ماشینی گروهی از الگوریتم‌های آموزش‌دیده را قادر می‌سازد تا پیش‌بینی‌هایی را براساس تشخیص الگو از مجموعه داده‌ای معین انجام دهند، که می‌تواند برای تصمیم‌گیری آگاهانه در هر حوزه‌ای، کسب‌وکار، سیاسی، اجتماعی، اقتصادی یا بهداشتی مورد استفاده قرار گیرد. نقش یک مهندس یادگیری ماشینی برای کسانی مناسب است که در تحقیقات کاربردی

و علم داده پیشینه دارند.

۳. علم داده:

هر کسب‌وکار یا سازمانی از دانشمندان داده، می‌خواهد که پیش‌بینی‌هایی را از داده‌ها به آن‌ها ارائه دهند که می‌تواند سود آن‌ها را به حداکثر برساند و فرآیندهایشان را ساده‌تر کند. علم داده تقریباً به تمام صنایعی که در حال حاضر مهم هستند مرتبط است. دانشمندان داده آگاه، ماهر و با استعداد در امور مالی، مراقبت‌های بهداشتی، فناوری و هر صنعت دیگری مورد نیاز هستند.

۴. تعامل انسان با کامپیوتر (HCI):

در واقع این علم نقطه تقاطع علوم رایانه و علوم رفتارشناسی، طراحی و چند علم دیگر است. ارتباط و تعامل انسان و رایانه از طریق واسطه اتفاق می‌افتد؛ که شامل نرم‌افزار و سخت‌افزار است. اگر بخواهیم تعریف دقیقی از این علم ارائه دهیم: علم تعامل انسان و رایانه یک رشته مرتبط با طراحی، ارزیابی و پیاده‌سازی سیستم‌های محاسباتی متقابل برای استفاده انسان در مطالعه پدیده‌های مهم پیرامون اوست. این رشته شاخه‌هایی از هر دو طرف درگیر را شامل می‌شود: مثلاً گرافیک کامپیوتری، سیستم‌عامل، زبان‌های برنامه‌نویسی، تئوری ارتباطات و طراحی صنعتی برای قسمت کامپیوتری؛ زبان‌شناسی، روانشناسی و کارایی انسان برای قسمت انسانی آن.

۵. امنیت سایبری:

امنیت سایبری روز به روز محبوب‌تر می‌شود زیرا کسب‌وکارهای بیش‌تر و بیش‌تری داده‌های حساس را به‌صورت آنلاین ذخیره می‌کنند، جدای از انجام تراکنش‌ها در حوزه دیجیتال؛ از تجارت الکترونیکی، بانکداری و



خدمات مالی گرفته تا هوافضا، موسسات دولتی، مخابرات و غیره، هر بخش که دارای ردپایی از دیجیتال است به یک متخصص امنیت سایبری نیاز دارد تا از سیستمها و دادههای حیاتی مرتبط با کاربر در برابر حمله و آسیب محافظت کند.

۶. پردازش ابری:

از آنجایی که کسبوکارهای بزرگ تا متوسط به ذخیرهسازی مقادیر زیادی از دادهها در انبارهای اینترنتی متصل میشوند که آنها توسط سرویس دهندههای بزرگ، مدیریت و معمولاً از آنها به عنوان «ابر» یاد می شود، نیاز به ذخیره، مدیریت و استفاده از دادهها دارند که این موضوع نیازمند مهارت‌های خاصی است که «رایانش ابری» از آن برخوردار است. کارشناسان رایانش ابری از تعداد زیادی خدمات پشتیبانی می کنند که شامل خدماتی مانند پشتیبان‌گیری ابری از GMAIL، عکس‌ها در تلفن هوشمند شما می شود، خدماتی که به شرکت‌های بزرگ اجازه می دهد همه دادههای خود را میزبانی کنند و همه برنامه‌های خود را در فضای ابری اجرا کنند.

نیاز به متخصصان رایانش ابری اکنون جهانی شده است، زیرا تعداد بیشتری از مشاغل و همچنین کاربران خردهفروشی از اینترنت برای ذخیره و دسترسی به دادهها در حال حرکت استفاده می کنند.

اگر زیاد اهل کدنویسی نیستید اما دوست دارید در بخش فناوری بمانید، این می تواند یک تخصص ایده آل برای شما باشد. این تخصص علوم کامپیوتر و کاربردهای تجاری آن را پوشش می دهد. عمدتاً بر همپستگی و تعامل بین هر دو مؤلفه تأکید دارد. جنبه‌های کلیدی آن عبارتند از: درک و طراحی فرآیندهای اطلاعاتی عمودی، ساختارهای داده، برنامه‌ها و سیستم‌هایی که نیازهای تجاری را برآورده می کنند.

۷. رسانه های دیجیتال/تعاملی:

اگر می خواهید وارد جنبه‌های خلاقانه شوید و همچنین می خواهید از دانش فنی و مدرک علوم کامپیوتر خود استفاده کنید، می توانید رسانه‌های دیجیتال، رسانه‌های تعاملی یا تخصص‌های مرتبط را انتخاب کنید. حوزه‌های کاربردی تحت پوشش این تخصص شامل بازی‌های رایانه‌ای، نرم افزارهای آموزشی، دستگاه‌های تلفن همراه، بازاریابی، تبلیغات و خدمات مبتنی بر مکان می باشد. صنعت رسانه‌های دیجیتال به سرعت در حال گسترش است و اکنون بخش مهمی از صنعت فناوری اطلاعات را تشکیل می دهد و این تخصص پاسخی به نیازهای این بخش مهم صنعت است.

همچنین اگر بخواهیم گرایش‌هایی برای مقطع دکتری رشته علوم کامپیوتر معرفی کنیم، می توان به موارد زیر اشاره کرد: علوم تصمیم و دانش، داده کاوی، منطق و روش های سوری، محاسبات نرم و هوش مصنوعی، نظریه سیستم‌ها، الگوریتم و نظریه محاسبه، محاسبات علمی.



یک مسئله حل نشده در علوم کامپیوتر

در رشته علوم کامپیوتر نیز مانند سایر رشته‌ها، مسئله حل نشده وجود دارند. در فهرست زیر با تعدادی از این مسئله‌ها آشنا می‌شویم:

۱- آیا تابع یک طرفه وجود دارد؟

۲- حدس بازی‌های یکتا.

۳- آیا فرضیه زمان اجرای نمایی درست است؟

۴- آیا می‌توان لگاریتم گسسته را در زمان اجرای چندجمله‌ای محاسبه کرد؟

۵- آیا می‌توان مسئله تجزیه اعداد را در زمان اجرای چندجمله‌ای بر روی یک رایانه عادی (غیر کوانتومی) حل کرد؟

البته مسئله‌های بسیاری وجود دارند که هنوز به صورت حل نشده باقی مانده‌اند.

اما از میان این مسئله‌ها، مسئله‌ای وجود دارد که رابطه عمیق و تنگاتنگی با ریاضیات دارد.

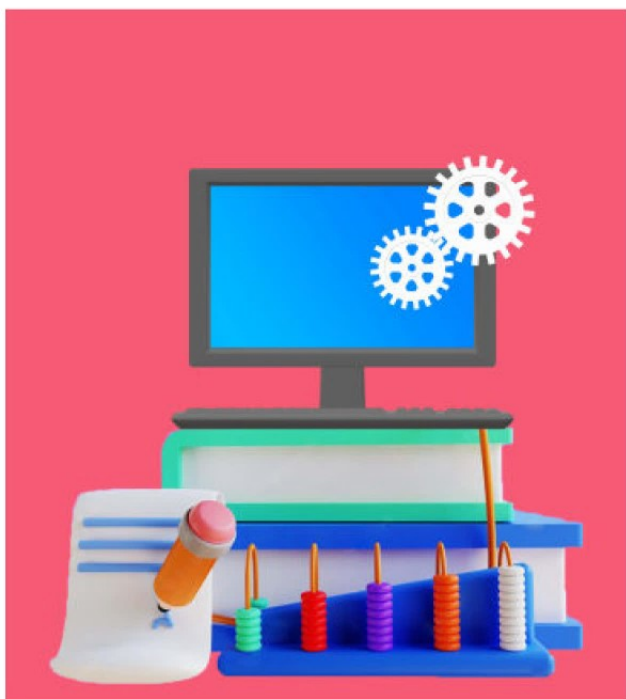
این مسئله بین ۷ مسائل جایزه هزاره ریاضیات وجود دارد. تمام این ۷ مسئله توسط انجمن ریاضی در سال ۲۰۰۰ مطرح شده‌اند. البته یکی از این مسائل در سال ۲۰۱۳ حل شد و اکنون از آن ۷ مسئله تنها ۶ مسئله بصورت حل نشده باقی‌مانده است. سوال این است که آیا برای تمام مسائلی

که الگوریتمی می‌تواند درستی یک جواب را بررسی کند؛ در زمان چندجمله‌ای، الگوریتمی وجود دارد که بتواند آن جواب را به همان سرعت پیدا کند؟ الگوریتم اول، دسته‌ای از پرسش‌ها را که NP نامیده می‌شوند توصیف می‌کند و الگوریتم دوم، دسته P را توصیف می‌کند. پرسش اصلی این است که آیا تمام پرسش‌هایی که در مجموعه NP هستند در مجموعه P هم هستند یا خیر؟ مسئله P و NP به‌طور عمومی یکی از مهم‌ترین سؤال‌های باز در ریاضیات و علوم نظری رایانه شناخته می‌شود؛ زیرا نتایج و تاثیرات زیادی بر روی مسائل دیگر در ریاضیات، زیست‌شناسی، فلسفه و رمزنگاری دارد.

«اسکات اندرسون» روزی درباره این مسئله چنین گفت: بیشتر ریاضیدانان و دانشمندان علوم کامپیوتر، انتظار دارند که $P \neq NP$ باشد. اگر $P=NP$ باشد، دنیا مکان کاملاً متفاوتی از چیزی که همیشه تصور می‌کردیم خواهد بود. زمانی که آن را پیدا کنیم، دیگر ارزش ویژه‌ای در جهش‌های خلاق یا اختلاف بنیادین بین حل کردن یک مسئله و پیدا کردن راه‌حل آن نخواهد بود. «مسئله‌ی NP در مقابل P همان مسئله‌ای است که از بین این ۷ مسئله، به علوم کامپیوتر مربوط است.»

از چرتکه تا کامپیوتر

نویسنده: کمران اعتمادمقدم



در این مقاله به بررسی سیر تحول کامپیوترها، از چرتکه تا کامپیوتر و همچنین بررسی کاربرد علم کامپیوتر در مدیریت صنایع می‌پردازیم. این حقیقت غیر قابل انکار است که در سده بیستام میلادی متخصصان موفق به ساخت یک دستگاه محاسبه‌گر چندمنظوره شدند و بعدها با افزودن قابلیت‌های مختلف آن را توسعه دادند. اما مبنا ساخت رایانه و شاید اصلی‌ترین دلیل ساخت رایانه‌ها نیاز بشر به انواع محاسبات بود. بشر اولیه از اعداد و علم حساب کاملاً کاربردی استفاده می‌کرد و استفاده از انگشتان برای محاسبه توسط بشر اولیه مبنایی برای به وجود آمدن نظام دستبندی ددهای بود؛ به مرور با استفاده بشر از وسایل گوناگون و مبنای دستبندی به وجود آمده منجر به اختراع چرتکه توسط چینی‌ها گردید. در آن زمان می‌توانستند چهار عمل اصلی را به سرعت انجام دهند و قدیمی‌ترین چرتکه گلی یافت شده مربوط به ۳ هزار سال پیش می‌باشد.

از اوایل قرن هفدهم میلادی تحولات تاریخی مختلفی رخ داد که بشر را قادر به ساخت ابزارهای محاسبه و کامپیوتر کرد و بعدها پیشرفت ایده‌ها و ماشین‌های گوناگون به ساخت ماشین‌ها و کامپیوترهای امروزی منجر شد. با این اوصاف ظهور کامپیوتر در حقیقت حاصل جستجو بشر برای دستیابی به دستگاه‌های محاسباتی سریع و دقیق بود که پیشرفت‌های سریع تکنولوژی در اوایل سده ۱۹ میلادی به دنبال داشت.

پس از چرتکه و ماشین حساب دستی پاسکال که به صورت مکانیکی به محاسبه جمع و تفریق می‌پرداخت؛ چارلز بابیج، پدر کامپیوترهای نوین در سده ۱۹ میلادی ماشینی را طراحی کرد که پیش‌آهنگ کامپیوترهای پیشرفته امروزی بود؛ نخستین اختراع آن «دستگاه تفاضلی» بود و ده سال بعد «دستگاه تحلیل» را ساخت که به کامپیوترهای امروزی شباهت

داشت.

در سال ۱۹۴۰ با بهبود تکنولوژی‌های برقی، پیشرفت‌های فنی، الکترونیکی و همچنین ساخت چند نوع ماشین محاسبه‌گر بزرگ منجر به ساخت کامپیوترهای بزرگی توسط دانشمندان مشهور آن زمان گردید.

هر بار با پیشرفت علوم مختلف نسل جدیدی از ماشین‌های محاسبه‌گر برای بهبود فرآیند و سرعت محاسبات به وجود آمدند و بعدها دانشمندان سعی کردند برای کامپیوترهایی که به وجود آمده در راستای بهبود موازی از «سرعت، نحوه ذخیره سازی، دقت، تعدد وظایف، خودکاری و...» بر روی این دستگاه‌ها داشته باشند.

دقیق به یک مدیر امکان جلوگیری از اقدامات نادرست را در بهترین زمان ممکن را می‌دهد. این اطلاعات از پردازش درست روی داده‌های خام به دست می‌آیند که منجر به داشتن یک «سیستم اطلاعاتی» خوب برای یک کسب‌وکار می‌شود.

- سیستم‌های انباری (انبارداری)، حسابداری، پرسنلی، پرداخت حقوق، کنترل پروژه و برنامه‌ریزی هم می‌توانند همانند مورد یک با بررسی و پردازش داده، یک «بانک اطلاعاتی» با توجه به ویژگی‌هایی مرتبط با عملکرد خود داشته و سرعت کارهای خود را بهبود دهند.
- همچنین کامپیوترها می‌توانند در مدیریت سیستم‌های تولیدی اتوماتیک بسیار مفید واقع شوند؛ در واقع سیستم‌های متعدد کنترل عددی و نوع کامپیوتری شده آن‌ها در صنایع مختلف کاربرد دارند و همچنین علاوه بر آن در طراحی و تولید کامپیوترها نیز به کمک صنایع مختلف می‌آیند.

در بررسی این مقاله ملاحظه کردیم که توسعه گسترده و سریع سخت‌افزار و نرم‌افزارهای کامپیوتری باعث گسترش و بهبود تکنولوژی مبتنی بر کامپیوترها گردید که همین امر باعث بهبود صنایع مختلف نیز می‌شود.

همچنین تلاش‌های گسترده که صورت گرفت باعث ایجاد نرم افزارهای مختلفی همچون «سیستم‌عامل‌ها، برنامه‌های مترجم، برنامه‌های از پیش ساخته یا از پیش نوشته، برنامه‌های کاربردی یا خاص استفاده‌کنندگان، برنامه‌سازی کامپیوتری، زبان‌های برنامه‌نویسی» و دیگر موارد از این قبیل گردید و باعث پیشرفت علم و کاربردی کردن این وسایل برای استفاده عمومیت جامعه شد.

همچنین در این مقاله پس از بررسی سیر تکامل کامپیوتر و ماشین‌های محاسبه‌گر به بررسی برخی از کاربردهای این ابزار در مدیریت صنایع پرداخته، که به معرفی کوتاهی از آنها می‌پردازیم:

- با توجه به افزونی داده‌ها همانند همه علوم این علم هم نیازمند سیستم‌های اطلاعاتی و داده‌پردازی است که به افراد و صنایع این امکان را می‌دهند تا اطلاعات مفید و کارآمد خود را تهیه و نگهداری کرده تا بتوانند به صورت قابل درک در کم‌ترین زمان در دسترس استفاده‌کنندگان قرار دهند. وجود اطلاعات صحیح، دسته‌بندی شده و

فشرده‌سازی کدگذاری هافمن

ذخیره سازی ۶۴ بیتی (با استفاده از رمزگذاری با طول ثابت) استفاده می‌کند. اگر دقت کنیم بسامد کاراکترهای $a \cdot b \cdot c$ و d به ترتیب ۴، ۲، ۱، ۱ است. حال بیا ببینیم سعی کنیم $aabacdab$ را با استفاده از تعداد بیت‌های کمتری با استفاده از این واقعیت که a بیشتر از b و b بیشتر از c و d تکرار شده است، نشان دهیم. با اختصاص تصادفی کد تک بیتی ۰ به a کد ۲ بیتی ۱۱ به b و کد ۳ بیتی ۱۰۰ و ۰۱۱ به ترتیب به کاراکترهای c و d شروع می‌کنیم.

```
a 0
b 11
c 100
d 011
```

بنابراین، رشته $aabacdab$ با استفاده از کدهای بالا به $00110100011011(11|0|011|100|0|11|0|0)$ کدگذاری می‌شود. اما مشکل واقعی در رمزگشایی نهفته است. اگر بخواهیم رشته 00110100011011 را رمزگشایی کنیم، به چند صورت می‌توان آن را رمزگشایی کرد که منجر به ابهام می‌شود.

```
0|011|0|100|011|0|11   adacdab
0|0|11|0|100|0|11|011  aabacabd
0|011|0|100|0|11|0|11  adacabab
...
and so on
```

کدگذاری هافمن (همچنین با نام رمزگذاری هافمن شناخته می‌شود) الگوریتمی برای فشرده‌سازی داده‌ها است و ایده اصلی فشرده‌سازی فایل را تشکیل می‌دهد. این مطلب درباره رمزگذاری با طول ثابت و طول متغیر، کدهای قابل رمزگشایی منحصر به فرد، قوانین پیشوند و ساخت درخت هافمن صحبت می‌کند.

ما از قبل می‌دانیم که هر نویسه (Character) دنباله‌ای از ۰ و ۱ است و با استفاده از ۸ بیت ذخیره می‌شود. این به عنوان «رمزگذاری با طول ثابت» شناخته می‌شود، زیرا هر نویسه از تعداد واحد ذخیره سازی بیت ثابت استفاده می‌کند.

با توجه به یک متن، چگونه می‌توان فضای مورد نیاز برای ذخیره یک نویسه را کاهش داد؟

سوال بالا همان ایده استفاده از «رمزگذاری با طول متغیر» است. ما می‌توانیم از این واقعیت استفاده کنیم که برخی از نویسه‌ها بیش‌تر از بقیه در یک متن تکرار می‌شوند، تا الگوریتمی طراحی کنیم که بتواند همان قطعه متن را با استفاده از تعداد بیت‌های کم‌تری نشان دهد. در رمزگذاری با طول متغیر، بسته به تکرار آن‌ها در متن داده شده، تعداد متغیری از بیت‌ها را به نویسه‌ها اختصاص می‌دهیم. بنابراین، برخی از نویسه‌ها ممکن است در نهایت یک بیت بگیرند، و برخی ممکن است در نهایت دو بیت بگیرند، برخی ممکن است با استفاده از سه بیت کدگذاری شوند، و غیره. مشکل رمزگذاری با طول متغیر در رمزگشایی آن نهفته است.

با توجه به دنباله‌ای از بیت‌ها، چگونه می‌توان آن را به طور یکتا رمزگشایی کرد؟ بیا ببینیم رشته $aabacdab$ را در نظر بگیریم. دارای ۸ کاراکتر است و از

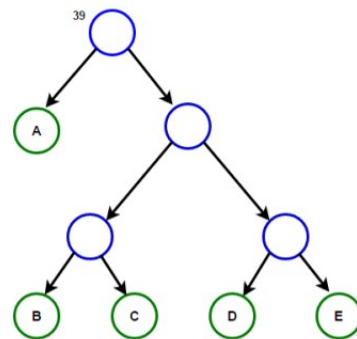
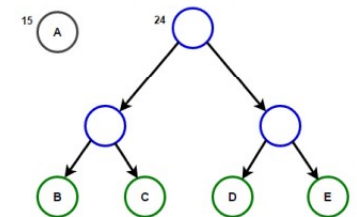
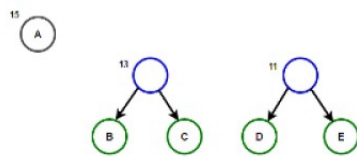
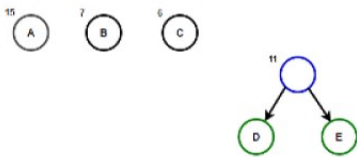
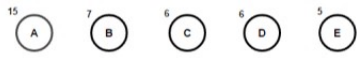
یک گره داخلی جدید با این دو گره به عنوان گره‌های فرزند و تکراری برابر با مجموع تکرارهای هر دو گره ایجاد کنید.

گره جدید را به صف اولویت اضافه کنید.

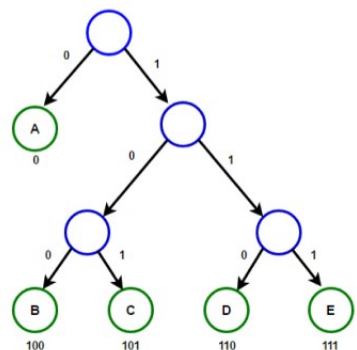
۳. گره باقی مانده گره ریشه است و درخت کامل است.

متنی را در نظر بگیرید که فقط از نویسه‌های «D»، «C»، «B»، «A» و «E» تشکیل شده است و تکرارهای آنها به ترتیب ۱۵، ۷، ۶، ۶، ۵ است.

شکل‌های صفحه بعد مراحل دنبال شده توسط الگوریتم را نشان می‌دهد.



مسیر از ریشه به هر گره برگ، کد پیشوند بهینه (که کد هافمن نیز نامیده می‌شود) مربوط به کاراکتر مربوط به آن گره برگ را ذخیره می‌کند.



برای جلوگیری از ابهامات در رمزگشایی، ما اطمینان حاصل می‌کنیم که رمزگذاری ما از «قاعده پیشوند» پیروی می‌کند، که منجر به «کدهای قابل رمزگشایی منحصر به فرد» می‌شود. قانون پیشوند بیان می‌کند که هیچ کدی پیشوند کد دیگری نیست. منظور ما از کد، بیت‌های استفاده شده برای یک نویسه خاص است. در مثال بالا، 0 پیشوند 011 است که قانون پیشوند را نقض می‌کند. اگر کدهای ما از قاعده پیشوند پیروی کنند، رمزگشایی بدون ابهام خواهد بود (و بالعکس).

بیباید مثال بالا را دوباره در نظر بگیریم. این بار کدهایی را به کاراکترهای 'a' و 'b' و 'c' و 'd' اختصاص می‌دهیم که قانون پیشوند را در نظر گرفته‌اند.

| | |
|---|-----|
| a | 0 |
| b | 10 |
| c | 110 |
| d | 111 |

با استفاده از کدهای فوق رشته aabacdab به 00100110111010. اکنون می‌توانیم 00100110111010 را به‌طور منحصر به فرد به رشته اصلی aabacdab رمزگشایی کنیم.

اکنون که در مورد رمزگذاری با طول متغیر و قانون پیشوند اطلاعاتی به‌دست آوردیم، اجازه دهید در مورد کدگذاری هافمن صحبت کنیم.

کد نویسی هافمن:

این تکنیک با ایجاد یک درخت دودویی از گره‌ها کار می‌کند. یک گره می‌تواند یک گره برگ یا یک گره داخلی باشد. در ابتدا همه گره‌ها، گره‌های برگ هستند که حاوی خود نویسه و وزن (تکرار) نویسه هستند. گره‌های داخلی حاوی وزن نویسه و پیوندهایی به دو گره فرزند هستند. به عنوان یک قرارداد رایج، بیت 0 نشان دهنده دنبال کردن فرزند چپ و بیت 1 نشان دهنده دنبال کردن فرزند سمت راست است. یک درخت تمام شده دارای n گره برگ و n-1 گره داخلی است. توصیه می‌شود که درخت هافمن باید نویسه‌های استفاده نشده را در متن کنار بگذارد تا بهینه‌ترین طول کد را تولید کند.

ما از یک صف اولویت برای ساخت درخت هافمن استفاده خواهیم کرد، جایی که گره با کمترین تکرار بیشترین اولویت را دارد. در زیر مراحل کامل آمده است:

۱. یک گره برگ برای هر نویسه ایجاد کنید و آنها را به صف اولویت اضافه کنید.

۲. در حالی که بیش از یک گره در صف وجود دارد:

دو گره با بالاترین اولویت (کمترین تکرار) را از صف حذف کنید.

بررسی کاربردهای زبان‌های برنامه‌نویسی مختلف



در این قسمت به بررسی کاربردهایی از برخی زبان‌های برنامه‌نویسی می‌پردازیم:

با توجه به این‌که تعدد زبان‌های برنامه‌نویسی بسیار است؛ برای دسته‌بندی بهتر زبان‌ها، آن‌ها را به چهار گروه زیر تقسیم‌بندی می‌کنیم:

۱. زبانهای برنامه نویسی ویندوزی
 ۲. زبانهای برنامه نویسی تحت وب
 ۳. زبانهای برنامه نویسی برای اپلیکیشن موبایل
 ۴. زبانهای برنامه نویسی چند منظوره
- پیش از آنکه بخواهید بهترین زبان برنامه نویسی را انتخاب کنید باید تصمیم بگیرید که به کدام حوزه علاقه دارید. در ادامه این مطلب ماسعی کردیم به بررسی کوتاهی از هر حوزه بپردازیم.
- زبان برنامه نویسی برای ویندوز:

به زبان‌هایی که برای ایجاد برنامه‌های تحت ویندوز مورد استفاده قرار می‌گیرند، زبان برنامه‌نویسی ویندوزی (desktop) گفته می‌شود. در واقع برنامه‌هایی که با استفاده از این زبان‌ها نوشته می‌شوند تنها روی سیستم‌عامل ویندوز قابل اجرا هستند. زبان‌های زیر پرکاربردترین زبان‌های این دسته هستند:

C, C++, C#, Java, Python, Delphi, Fox Pro, Visual Basic, Perl, Ruby, GO/Golang

- زبان برنامه نویسی تحت وب:

اگر بخواهید برای مرورگرها برنامه‌ای طراحی کنید (مثل فرم‌های ثبت‌نام برخط) باید یکی از زبان‌های برنامه نویسی تحت وب را یاد بگیرید. البته باید بدانید زبان‌های تحت وب هم به دو دسته‌ی سمت کاربر (Front-End)

و سمت سرویس‌دهنده (سرور) (Back-End) تقسیم می‌شوند. زبان‌های برنامه نویسی سمت کاربر:

CSS, HTML, jQuery, JavaScript

زبانهای برنامه نویسی سمت سرویس‌دهنده:

ASP.NET, PHP, Ruby, Java, Python, Golang

- زبان برنامه نویسی موبایلی کدامند؟

اگر دوست دارید برای موبایل تحت سیستم‌عامل اندروید یا IOS، اپلیکیشن طراحی کنید، پس باید یکی از زبان‌های زیر را یاد بگیرید.

C#, Python, Java, JavaScript, Perl, Ruby, Swift, Objective-C

زبان برنامه نویسی چند منظوره:

اگر به لیست زبان‌هایی که تا الان نام بردیم دقت کرده باشید، متوجه تکرار چند اسم خواهید شد. به این‌ها زبان‌های برنامه نویسی چندمنظوره (General-Purpose Language) گفته می‌شود. یعنی شما با یادگیری یکی از این زبان‌ها می‌توانید هم برای بسترهای مختلف «ویندوز، وب، اندروید و IOS» برنامه طراحی کنید. زبان‌های زیر در این دسته قرار می‌گیرند.

C, C++, C#, Java, Python, Ruby, GO, Scala

• پایتون، مار محبوب دنیای برنامه نویسی

در دنیای برنامه نویسی، پایتون یک زبان چندمنظوره، کاربردی، سطح بالا (High Level) و در عین حال ساده است. به همین خاطر برای خیلی از تازه‌واردان دنیای برنامه‌نویسی پایتون زبان جذابی است. چون با یادگیری این زبان خیلی زود می‌توانند کدنویسی را شروع کنند. بسیاری از شرکت‌های بزرگ مثل: گوگل، یوتیوب، فیسبوک، ناسا، یاهو، دراپ باکس و بسیاری دیگر برای توسعه سایت‌شان از زبان اسکریپتی و متن‌باز پایتون استفاده می‌کنند. پایتون یک زبان شیء‌گرا است که در زمینه‌های متنوعی مثل «تحلیل داده‌ها، توسعه هوش مصنوعی، یادگیری ماشینی، توسعه وب، خودکارسازی سامانه، هک، واسط‌های برنامه‌نویسی و ...» کاربرد دارد.

• جاوا، یک فنجان قهوه با طعم کد:

فرزند خانواده‌ی C پا به عرصه‌ی IT گذاشت و در سال ۲۰۰۰ هم با تولد #C، این خانواده نوهدار شد.

زبان C یک زبان چندمنظوره و سطح میانی است. سی، زبان کوچکی است و کلمات کلیدی محدودی دارد. از طرف دیگر زبانی ساخت‌یافته، دستوری، منعطف و قابل حمل است. باید بدانید مهم‌ترین ویژگی‌هایی که زبان ++C را از C متمایز می‌کند، یکی ساختار شی‌گرایی و دیگری استفاده از کلاس است. علاوه بر این، زبان ++C بیش‌تر نزدیک به زبان سطح بالاست. با سی شارپ، نوه‌ی خانواده‌ی سی آشنا شوید #C یک زبان چندمنظوره و شی‌گراست. اولین بار شرکت مایکروسافت این زبان کاربردی را وارد دنیای برنامه‌نویس‌ها کرد. زبانی که سال به سال و نسخه به نسخه بروز می‌شود و قابلیت‌های جدیدی پیدا می‌کند. بی دلیل نیست که در فهرست انجمن TIOBE، زبان سی شارپ توانسته در بین این همه زبان برنامه نویسی رتبه چهارم را به دست بیاورد. #C یک زبان سطح بالاست که برای شروع آموزش برنامه نویسی سی شارپ انتخاب بسیار مناسبی است. چون زبانی قدرتمند، فراگیر و کاربردی است که یادگیری‌اش نسبتاً آسان است. به همین خاطر می‌توانید به کمکش خیلی زود کدنویسی را شروع کنید.

دنیای برنامه‌نویسی و زبان‌های آن بسیار گسترده است، در این بخش سعی کردیم به مشهورترین موارد بپردازیم اما شما می‌توانید برای دریافت اطلاعات بیش‌تر به وب‌گاه‌های مختلف مراجعه و اطلاعات مفیدی کسب کنید.



مهارت‌های مورد نیاز

در علوم کامپیوتر

در این بخش می‌خواهیم مهارت‌های لازم برای یادگیری رشته‌ی علوم کامپیوتر و همچنین چگونگی ارتقا و یادگیری این مهارت‌ها را مورد بررسی قرار دهیم. کسانی که در این رشته در حال تحصیل هستند برای موفقیت در رشته‌ی خود باید طیفی از مهارت‌های تکنیکی و نرم را یاد بگیرند. از آن جایی که این رشته، رشته‌ای بسیار گسترده بوده و تمرکز وسیع تخصصی فراوانی دارد انتخاب و یادگیری مهارت به علاقه و ماهیت خود شخص و رشته انتخابی بسیار وابسته است.

در این جا به بررسی چند مهارت سخت (مانند کد نویسی توسعه ی نرم افزار و نگارش فنی) و مهارت‌های نرم (مانند ارتباطات و کار تیمی) می‌پردازیم.

• مهارت‌های تکنیکی و ریاضی:

همان‌طور که در ابتدا گفته شد، مهارت‌های فنی این رشته بسیار گسترده می‌باشد و متخصصان را قادر می‌سازد در زمینه‌های توسعه‌ی نرم‌افزار مهندسی سخت افزار و برنامه نویسی مهارت داشته باشند همچنین باید مهارت‌هایی در زمینه ریاضیات شامل آمار، جبر، ریاضیات گسسته و دیفرانسیل داشته باشند.

• مهارت‌های نوشتاری فنی:

همه متخصصان در این رشته از این مهارت استفاده نمی‌کنند ولی یادگیری آن بسیار مهم است. زیرا با دانستن این مهارت این اجازه به متخصصان داده می‌شود که اطلاعات فنی را به شیوه‌ای منتقل کنند که برای همه قابل درک باشد.

• مهارت مدیریت پروژه:

فارغ از تخصص فنی آن‌ها (مانند برنامه نویسی و توسعه نرم افزار) دانشمندان رایانه ممکن است مهارت‌های مدیریت پروژه را برای رهبری و مدیریت تیم‌های فناوری اطلاعات هنگام تکمیل پروژه‌های مختلف به کار ببرند.

توانایی استراتژی کنترل بودجه، پیش‌بینی نتایج کاهش خطرات، سایر فرایندهای برنامه نویسی توسعه و شروع یک پروژه هر کدام جز مهارت‌هایی هستند که متخصصان این رشته باید برای موفقیت در نقش خود به آن تکیه کنند.

• مهارت‌های تحلیلی:

دو ویژگی که متخصصان در کار از آن‌ها استفاده می‌کنند به ترتیب تحلیل داده‌ها و آزمایش کاربردهای فنی می‌باشد. در همه‌ی حوزه‌ها مثل آزمایش، جمع آوری، ارزیابی و مستند سازی داده‌ها و حتی در استفاده از تکنیک‌های مختلف در محل کار به طور منظم از توانایی تحلیلی خود استفاده می‌کند.

• مهارت حل مسئله:

حل مسئله یکی از مهارت‌های نرم به شمار می‌رود. حل مسئله از رویکردهای سیستماتیک و منطقی در این رشته هست که دانشمندان برای تکمیل پروژه‌های پیچیده و دستیابی به اهداف تجاری به آن نیاز دارند.



- **مهارت تفکری انتقادی:**

این مهارت شامل ارزیابی و ترکیب اطلاعات برای رسیدن به نتیجه ای مطلوب و سودمند می‌باشد. از آنجایی که متخصصان این رشته با فناوری‌های پیچیده سروکار دارند باید این مهارت را در خود تقویت کنند. با تقویت این مهارت می‌توانند رویکردهای متناسب با پروژه را انجام دهند و اغلب زمان را برای تکمیل اهداف فنی مهم کاهش دهند.

- **خلاقیت:**

در این رشته مقدار زیادی خلاقیت وجود دارد. برای توسعه برنامه‌ها و برنامه‌های کاربری جدید و تکمیل پروژه‌های فنی نیاز به خلاقیت و نوآوری دارند.

- **تیم گروهی و مهارت‌های بین فردی:**

این مهارت جز مهارت‌های نرم این رشته محسوب می‌شود. اغلب متخصصان برای این که در تیم‌های مربوط به پروژه‌های فنی مشغول

به کار شوند باید مهارت‌هایی مانند مهارت‌های بین فردی و کار گروهی برای هدایت موفقیت آمیز این محیط استفاده کنند. مهارت‌هایی که اغلب باید متخصصان استفاده کنند: توانایی تفویض اختیار، هدایت، همکاری، به اشتراک گذاشتن ایده‌ها، درک دیدگاه‌ها و ...

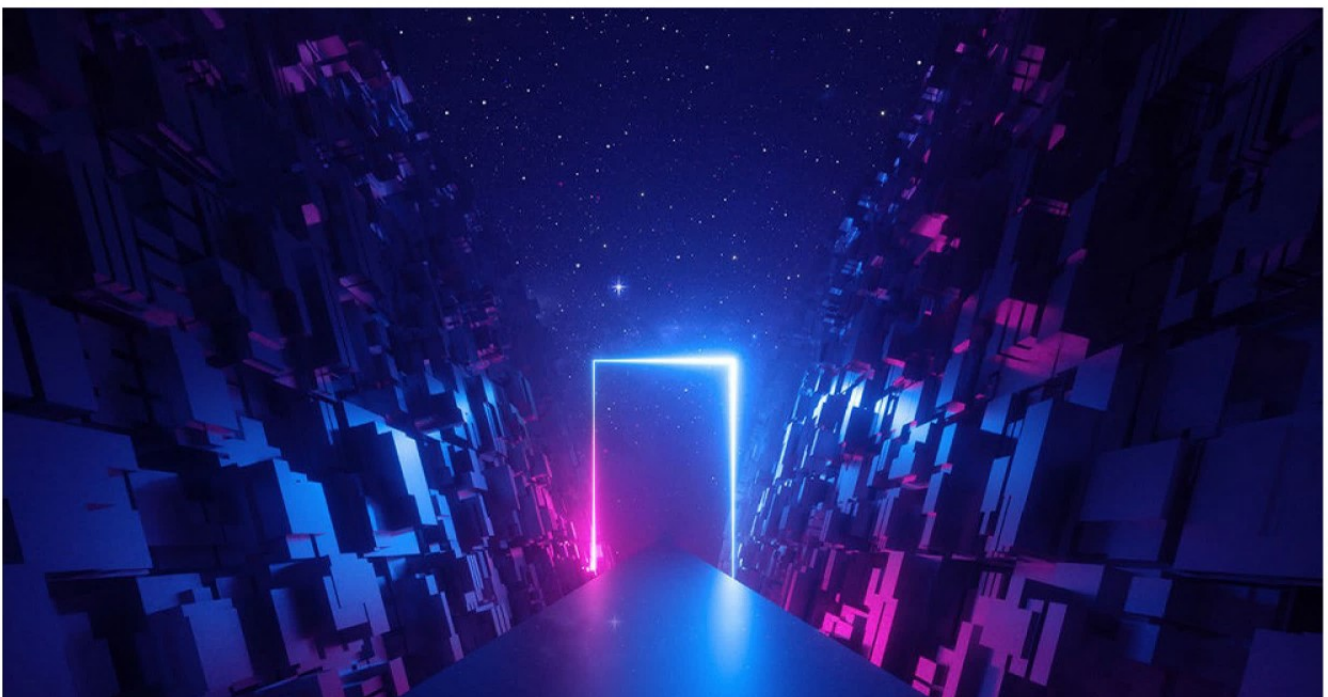
- **مهارت‌های ارتباطی:**

به‌طور کلی این مهارت در هر رشته و کاری ضروری و امری واجب است. با این حال در رشته علوم کامپیوتر نیاز بیشتری به این مهارت دیده می‌شود. به‌طور مثال وقتی متخصصان روی یک پروژه کار می‌کنند از مهارت‌های نوشتاری برای انتقال اطلاعات استفاده می‌کنند در کنار این مهارت‌های ارتباط کلامی و غیر کلامی (گوش دادن فعال، پرسیدن سوال، استفاده از ابزارهای ارتباطی) از مهم‌ترین مهارت‌ها برای متخصصان این رشته خواهد بود.

متاورس، نسل بعدی اینترنت

بسیاری از مفاهیم فناورانه دیگر، در ابتدا در حد یک رؤیای پردازی در ذهن چند نویسنده و فیلم‌ساز بوده که اخیراً به مرحله اجرا رسیده است؛ در حقیقت واژه متاورس و خلق آن به دهه ۹۰ میلادی زمانی که «نیل استفنسون» کتابی با نام (Snow Crash) منتشر کرد باز می‌گردد. متاورس، یک جهان مجازی است که کاربران می‌توانند با استفاده از عینک‌های مخصوص به نام عینک واقعیت مجازی شخصیت موردعلاقه خود را در آن بسازند و به‌واسطه چهرک (آواتار، تصویر مجازی و شخصیتی که از خود می‌سازند)، درست مانند زندگی واقعی، در آنجا زندگی کنند. شرکت‌هایی که در زمینه نوآوری تکنولوژی‌های جدید کار می‌کنند، از

در این قسمت می‌خواهیم به بررسی یک خبر جدید در حوزه فناوری پردازیم؛ نامی که این روزها خیلی می‌شنویم، «متاورس» است. متاورس از دو واژه (Meta) یعنی جهانی و (Verse) یعنی برتر تشکیل شده است. در نتیجه عبارت متاورس به معنای «جهان برتر» یا «برتر از جهان» است؛ دنیایی که محدودیت‌های دنیای فیزیکی را پشت سر می‌گذارد و شاید به نسخه‌ای از آرمان‌شهر موردعلاقه انسان‌ها باشد. متاورس پدیده‌ای نوظهور است که با وجود بیش از چند دهه قدمت این واژه، به‌تازگی مفهومی عملیاتی به خود گرفته و این روزها بسیاری از علاقه‌مندان به تکنولوژی را سرگرم خود کرده است. این تکنولوژی شبیه





آواتار در جلسات اداری مجازی برای سر و صورت دادن به کارهای استفاده می‌کند. این شخص پس از پایان جلسه، بلیت یک کنسرت مجازی را می‌خرد و همراه با چهره‌های دوستانش در میان صدها شرکت کننده در کنسرت ظاهر می‌شود. هنگامی که کنسرت موسیقی تمام می‌شود، گروه به او یادآوری می‌کنند که خریدن تی‌شرت را فراموش نکند. او طرح‌هایی را که امروز در غرفه‌های فروشگاه‌های اینترنتی دیده، مرور می‌کند. سپس برای خرید یکی از آنها با ارز دیجیتال پرداخت می‌کند و روز بعد آن را در دفتر کار مجازی می‌پوشد.

یکی از همکاری‌های او خواهد پیراهن او را برای فرزندش قرض بگیرد تا در آن شب در یک بازی Roblox بپوشد و او پیراهن را به همکاری‌اش قرض می‌دهد.

در این سناریوی ساده که مثال زده شد، بسیاری از ابزارهای ارتباطی شرکتی، پخش زنده رویدادها، تجارت الکترونیک و اطلاعات ارزشمند اشتراک گذاری شده است. این تعاملات تنها در صورتی کار می‌کند که هر شرکت ارائه دهنده امکانات، سیستم خود را به گونه‌ای بسازد که دارایی‌هایی مانند آواتارها و یا پیراهن‌ها با همدیگر سازگار بوده و قابل انتقال باشند.

فناوری متاورس جهت ارائه تجربه جدید و پیشرفته‌ای برای تقویت و توسعه دنیای سه بعدی استفاده می‌نماید. کلمه متاورس یک دنیای دیجیتالی را توصیف می‌کند که مشابه یک دنیای واقعی کاملاً تحقق یافته و فراتر از دنیایی است که در آن زندگی می‌کنیم.

در حال حاضر متاورس به عنوان یک جهان مجازی قلمداد می‌شود که تمامی جنبه‌های فناوری‌های دیجیتال شامل ویدئو کنفرانس، بازی‌های ویدئویی اینترنتی مانند ماین کرافت یا Roblox، معاملات ارزهای دیجیتال، ارسال و دریافت ایمیل، واقعیت مجازی، رسانه‌های اجتماعی و پخش‌های زنده را با هم ترکیب می‌کند.

فعالیت‌های متخصصان بر روی اینکه چگونه می‌توان این بخش‌های مجازی را با هم هماهنگ نمود، در حال پیشرفت است. اما بعضی از بزرگان دنیای فناوری، متاورس را آینده ارتباطات و تعامل انسانی می‌دانند.

تفاوت دنیای آنلاین با متاورس چیه؟

دنیای متاورس اختلاطی از کار و بازی به وجود می‌آورد. به عنوان مثال، کاربری یک چهره سه بعدی که در واقع یک نمایش دیجیتالی از خود او می‌باشد، در فیس بوک یا تیم‌های مایکروسافت ایجاد می‌کند. سپس از آن

معرفی کتاب و دوره

در علوم کامپیوتر

کتاب:

کد: زبان پنهان سخت افزار و نرم افزار کامپیوتر

کدها، روش‌های مبتکرانه‌ای را به ما نشان می‌دهند که ما زبان را دستکاری می‌کنیم و ابزار جدیدی برای برقراری ارتباط با یکدیگر اختراع می‌کنیم. و از طریق کد، می‌بینیم که چگونه این نبوغ و اجبار بسیار انسانی ما برای برقراری ارتباط، نوآوری‌های تکنولوژیکی دو قرن گذشته را هدایت کرده است.

نویسنده چارلز پترولد با استفاده از اشیاء روزمره و سیستم‌های زبانی آشنا مانند بریل و کد مورس، روایتی روشن‌گر را برای هر کسی که تا به حال در مورد زندگی درونی پنهان رایانه‌ها و سایر ماشین‌های هوشمند فکر کرده است، می‌سازد.

کد تمیز: کتاب راهنمای برای ساخت نرم افزاری Agile

حتی کد بد هم می‌تواند کار کند. اما اگر کد تمیز نباشد، می‌تواند یک سازمان توسعه را به زانو درآورد. هر سال، ساعت‌های بی شماری و منابع قابل توجهی به دلیل کد نوشته شده ضعیف از بین می‌روند. اما لازم نیست اینطور باشد.

رابرت سی. مارتین، کارشناس مشهور نرم افزار، الگوی انقلابی را با کتاب کد تمیز ارائه می‌دهد: کتابچه راهنما برای ساخت نرم افزار چابک. مارتین با همکاری از Object Mentor بهترین روش چابک خود را در پاکسازی کد به شکل کتابی قابل فهم در آوردند که شما را به برنامه‌نویس بهتری تبدیل می‌کند.

در بخش‌های دیگر نشریه بیشتر با این کتاب آشنا خواهید شد.

الگوریتم:

الگوریتم‌های رابرت سجویک و کوین وین، کتاب درسی پیشرو در مورد الگوریتم‌های امروزی است و به طور گسترده در دانشکده‌ها و دانشگاه‌های سراسر جهان استفاده می‌شود. این کتاب مهم‌ترین الگوریتم‌های رایانه‌ای را بررسی می‌کند که در حال حاضر مورد استفاده قرار می‌گیرند و به بررسی کامل ساختارهای داده و الگوریتم‌ها برای مرتب‌سازی، جستجو، پردازش نمودار و پردازش رشته‌ها می‌پردازد که شامل پنجاه الگوریتمی است که هر برنامه‌نویسی باید بداند. الگوریتم‌های این کتاب نشان دهنده مجموعه‌ای از دانش است که طی ۵۰ سال گذشته توسعه یافته است و نه تنها برای برنامه‌نویسان حرفه‌ای و دانشجویان علوم کامپیوتر بلکه برای هر دانش‌آموزی که علاقه‌مند به علوم، ریاضیات و مهندسی است، غیر قابل ذکر است.

دوره:

۱. قطعا یکی از بهترین سایت‌ها برای یادگیری برنامه‌نویسی و مهارت‌های مرتبط به حوزه علوم کامپیوتر سایت [codecademy](https://www.codecademy.com) است که منابع آموزشی متنوعی را به تفکیک برای یادگیری افراد علاقه‌مند در سایت خود قرار داده است.

۲. سایت [W3Schools](https://www.w3schools.com) هم یکی دیگر از مهم‌ترین سایت‌ها در زمینه یادگیری مهارت‌های برنامه‌نویسی می‌باشد که در این سایت هم به تفکیک مطالب منابع و تمرین متنوعی را در حوزه‌های مختلف قرار داده است.

انتخاب رشته مناسب علوم کامپیوتر

برای کارشناسی ارشد

برای انتخاب رشته مناسب در علوم کامپیوتر:

۱. زمینه های مورد علاقه خود را پیدا کنید.

۲. علایق خود را با نکات مثبت خود مانند پیشینه فنی، پروژه های دانشگاهی، کارآموزی، تجربه کاری هماهنگ کنید.

۳. اهداف شغلی خود را چه کوتاه مدت و چه بلند مدت مشخص کنید.

حالا:

• اگر به دنبال دیدگاه دانش گسترده ای از علوم کامپیوتر هستید، پس یک کارشناسی ارشد عمومی در علوم کامپیوتر مناسب است.

• اگر می خواهید در یک زمینه تخصص داشته باشید، می توانید توانایی خود را بهبود ببخشید.

• اگر عاشق کدنویسی هستید، برنامه نویسی نرم افزار گزینه ای عالی برای شما به عنوان یک تازه کار است.

• اگر تجربه کاری همراه با هوش تجاری دارید، می توانید علم داده یا تجزیه و تحلیل داده را انتخاب کنید.

• جدا از کدنویسی قوی، اگر تسلط استثنایی بر آمار ریاضی و عشق دارید، می توانید هوش مصنوعی را در نظر بگیرید.

• اگر دوست دارید داده ها را تجزیه و تحلیل کنید، الگوها را شناسایی کنید، و بینش هایی را به دست آورید که منجر به تصمیم گیری های تجاری آگاهانه تر می شود، می توانید هر بخشی را از میان هوش مصنوعی، علم داده، تجزیه و تحلیل داده ها یا نظریه رایانه انتخاب کنید. توجه داشته باشید که این نقش ها همچنین مستلزم ارائه نتیجه تجزیه و تحلیل داده ها به ذینفعان کلیدی به زبان واضح است.

• اگر به طراحی بصری توجه دارید و به جزئیات توجه می کنید، پس نقشی را به عنوان یک برنامه نویس، طراح بازی، طراح گرافیک کامپیوتری یا طراح رابط کاربری در نظر بگیرید. این نقش ها همچنین نیازمند خلاقیت و انعطاف پذیری برای برآوردن نیازهای ذینفعان داخلی و خارجی است.

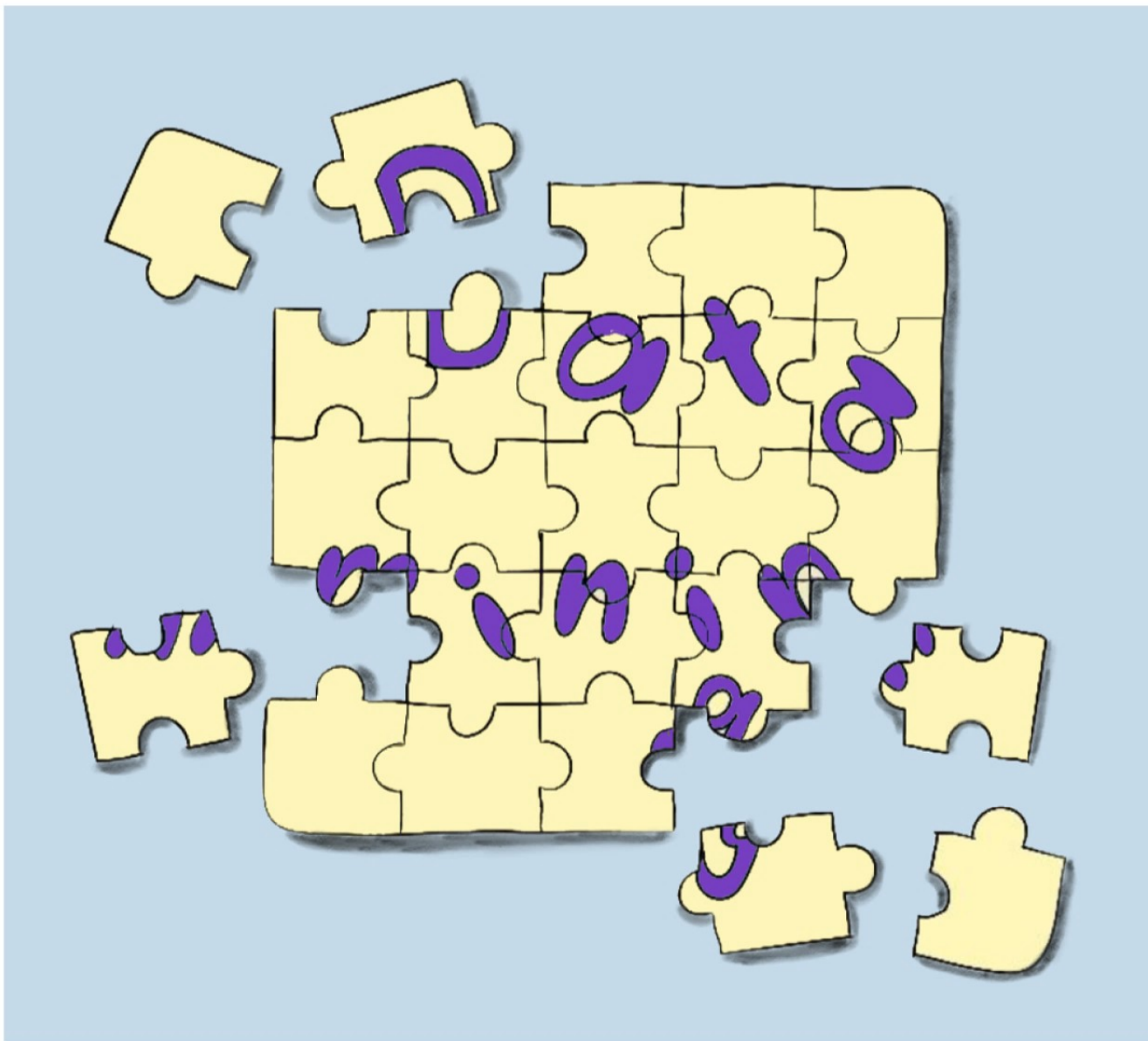
• اگر سابقه آموزش فنی و عملیات پشتیبان دارید، نقشی را به عنوان یک مهندس، مدیر سیستم، یا سرپرست شبکه در نظر بگیرید.

• اگر دارای تخصص فنی و همچنین توانایی مذاکره دیپلماتیک هستید، نقشی را در امنیت اطلاعات در نظر بگیرید. نقش احتمالی دیگر، مهندس DevOps - مخفف «عملیات توسعه دهنده» است، نقشی که به سازمان ها کمک می کند شکاف بین سرعت و ثبات را در چرخه عمر توسعه پر کنند.

داده‌کاوی



0010



داده‌کاوی،

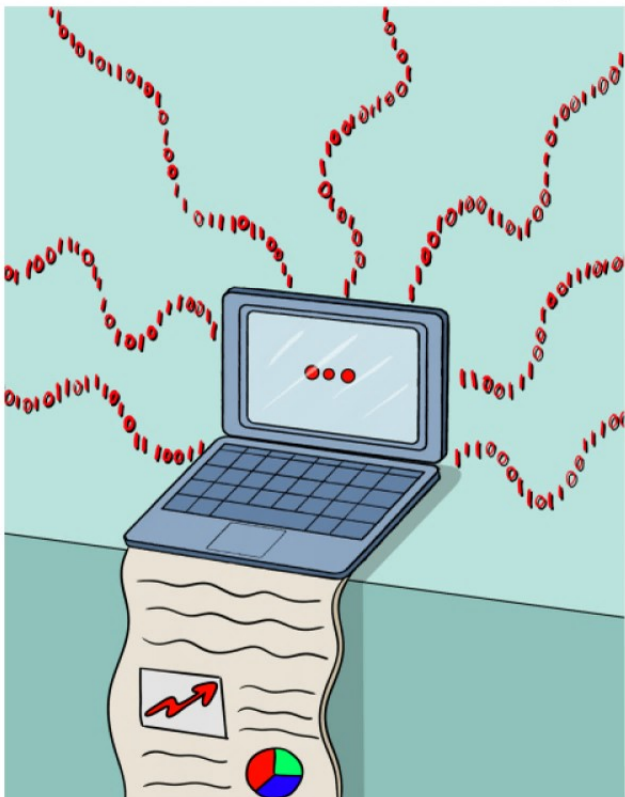
کشف دانش در پایگاه‌داده‌ها

داده‌کاوی فرآیندی است که در آغاز دهه ۹۰ مطرح شد و با نگرشی نو، به مسئله استخراج اطلاعات از پایگاه داده‌ها می‌پردازد. از سال ۱۹۹۵ داده‌کاوی به صورت جدی وارد مباحث آمار شد و در سال ۱۹۹۶، اولین شماره مجله کشف دانش و معرفت از پایگاه داده‌ها منتشر شد. محققانی نظیر «براجمن و آناند» (۱۹۹۶) کلیه مراحل واقع‌گرایانه و رو به جلو کشف دانش از پایگاه داده‌ها را تشخیص دادند.

در سال ۱۹۸۹ و ۱۹۹۱ کارگاه‌های کشف دانش از پایگاه داده‌ها توسط «پیاتسکی» و همکارانش برگزار شد. در فواصل سال‌های ۱۹۹۱ تا ۱۹۹۴ کارگاه‌های کشف دانش از پایگاه داده‌ها توسط «فیباد و پیاتسکی» و دیگران برگزار شد. به طور رسمی اصطلاح داده‌کاوی برای اولین بار توسط «فیباد» در اولین کنفرانس بین‌المللی «کشف دانش و داده‌کاوی» در سال ۱۹۹۵ مطرح شد و این اصطلاح در جامعه هوش مصنوعی و یادگیری ماشین محبوب شد و داده‌کاوی در جوامع تجاری و مطبوعاتی نفوذ بیشتری پیدا کرد. همچنین در جامعه دانشگاهی، انجمن‌های اصلی برای تحقیق در

مهندسی و کسب و کار شده است. زمینه داده‌کاوی و کشف دانش از پایگاه داده‌ها به عنوان یک رشته علمی جدید در مهندسی و علوم کامپیوتر ظهور کرده است. همچنین مهندسی صنایع با حوزه‌های گوناگون و دربر داشتن فرصت‌های بی‌نظیر اکنون برای کاربرد داده‌کاوی و کشف دانش از پایگاه داده‌ها و برای توسعه مفاهیم و روش‌های تازه در این زمینه آماده است؛ چرا که فرآیندهای صنعتی زیادی اکنون برای مطمئن شدن از کیفیت سفارشات محصول و کاهش هزینه‌های محصول به طور خودکار و کامپیوتری شده‌اند. انواع تجزیه و تحلیل در داده‌کاوی: به طور کلی انواع تجزیه و تحلیل در داده‌کاوی را می‌توان با دسته‌های زیر توصیف کرد:

۱. تجزیه و تحلیل توصیفی: این بخش به پاسخ سوالات در مورد آنچه اتفاق افتاده است کمک می‌کند.
۲. تجزیه و تحلیل تشخیصی: این بخش به پاسخ سوالات در مورد چرایی اتفاقات کمک می‌کند. این تکنیک‌ها تجزیه و تحلیل‌های توصیفی اولیه را تکمیل‌تر می‌کنند.
۳. تجزیه و تحلیل پیش‌گویانه: این قسمت به پاسخگویی به سوالاتی در مورد آنچه در آینده رخ خواهد داد کمک می‌کند. این قسمت یکی از بخش‌های مهم است که تکنیک‌های آن شامل انواع تکنیک‌های آماری و یادگیری ماشینی مانند درخت‌های تصمیم‌گیری و سرفیت می‌باشد.
۴. تجزیه و تحلیل تجویزی: در این بخش با استفاده از بینش‌های حاصل از تجزیه و تحلیل پیش‌بینی، می‌توان تصمیمات مبتنی بر داده را اتخاذ کرد.



سال ۱۹۹۵ شروع به کار کردند، زمانی که همین کنفرانس بین‌المللی در مونترال تحت حمایت AAAI برگزار شد. امروزه کنفرانس‌های مختلفی در این زمینه در سراسر دنیا برگزار می‌شود.

شرکت‌ها به واسطه خدماتی که به مشتریان عرضه می‌کنند با آن‌ها ارتباط مستمر دارند؛ از سوی دیگر اطلاعات زیادی از مشتریان به دست می‌آوردند و این موضوع سبب می‌شود که داده‌های زیادی برای تحلیل رفتار مشتری در دست داشته باشند که تحلیل درست آن‌ها، شرکت را به سودآوری بیشتر می‌رساند.

داده‌کاوی در واقع یک روش است، روشی که به واسطه آن مسئله‌ای حل می‌شود. این روش بر روی حجم زیادی از داده‌ها کار می‌کند و تحلیل‌های مورد نیاز را انجام می‌دهد، در نهایت یک سری الگوهای تکرارشونده استخراج می‌شود که می‌تواند برگ برنده‌ای برای آن شرکت یا موسسه باشد. سپس نوبت به پیدا کردن ارتباطات بین الگوهای مختلف است و در نهایت یک سری راه‌حل‌های مهم در این چالش وارد می‌شوند.

داده‌کاوی بر روی یک سری اطلاعات کار می‌کند که در نگاه اول احتمالاً کاربرد زیادی نخواهند داشت اما وقتی پژوهش‌های دقیقی صورت می‌گیرد نتایج به دست آمده بسیار ارزشمند می‌شود.

دنیای امروز، دنیای ارتباطات است، ارتباط بین مکان و زمان‌های مختلف به طوری که اغلب موارد در بستر مجازی دنبال می‌شود. اطلاعاتی که قرار است از مشتریانی جمع‌آوری شود که شاید هرگز دیده نشوند، یک موهبت بزرگ برای شرکت‌ها محسوب می‌شود و اهمیت زیادی دارد.

این علم در شرکت‌های بزرگ آن‌قدر مهم و شناخته شده است که بسیاری از برنامه‌ریزی‌ها بر اساس آن صورت می‌گیرد. به عنوان مثال، پیش از طراحی یک پروژه پرهزینه نیاز است که اطلاعات زیادی در مورد این پروژه مورد بررسی قرار گیرد تا بهترین تصمیم برای نحوه اجرای پروژه یا عدم اجرای آن گرفته شود.

در حال حاضر، داده‌کاوی مهم‌ترین فناوری جهت بهره‌برداری موثر از داده‌های حجیم است و اهمیت آن رو به فزونی می‌باشد. به طوری که تخمین زده شده است که مقدار داده‌ها در جهان هر ۲۰ ماه به حدود دو برابر می‌رسد. در یک تحقیق که بر روی گروه‌های تجاری بزرگ در جمع‌آوری داده‌ها صورت گرفت مشخص گردید که ۱۹ درصد از این گروه‌ها دارای پایگاه داده‌هایی با سطح بیشتر از ۵۰ گیگابایت می‌باشند و ۵۹ درصد از آن‌ها انتظار دارند که در آینده‌ای نزدیک در چنین سطحی قرار گیرند در صنایعی مانند کارت‌های اعتباری و ارتباطات و فروشگاه‌های زنجیره‌ای و خریدهای الکترونیکی و اسکنرهای بارکد خوان هر روزه داده‌های زیادی تولید و ذخیره می‌شوند.

افزایش داده‌های بسیار باعث پیدایش فرصت‌های تازه برای کار در علوم

کاربردهای داده‌کاوی

۲- تحلیل سبد بازار (خرده فروشی‌ها): فروشگاه‌های مواد غذایی، از کاربران شناخته‌شده‌ی تکنیک‌های داده‌کاوی هستند. بسیاری از سوپرمارکت‌ها کارت‌های وفاداری رایگان را به مشتریان ارائه می‌دهند که به آنها امکان دسترسی به قیمت‌های کاهش‌یافته را می‌دهد که برای افراد غیرعضو در دسترس نیست. این کارت‌ها به فروشگاه‌ها این امکان را می‌دهد که ردیابی کنند که چه کسی، چه زمانی و به چه قیمتی محصولی را خریداری کرده‌است. پس از تجزیه و تحلیل داده‌ها، فروشگاه‌ها می‌توانند از این داده‌ها برای ارائه کوپن‌هایی شخصی‌سازی شده با عادت‌های خرید مشتریان به آنها ارائه دهند و تصمیم بگیرند که چه زمانی اقلام را در معرض فروش قرار دهند و یا چه زمانی آنها را به قیمت کامل بفروشند.

۳- سرگرمی: سرویس‌های رسانه جاری مثل نتفلیکس یا اسپاتیفای و... داده‌کاوی می‌کنند تا آنچه را که کاربران تماشا می‌کنند یا گوش می‌دهند را شناسایی کرده و توصیه‌های شخصی‌سازی شده براساس عادات تماشا و شنیدن افراد به آنها ارائه دهند.

۴- خدمات مالی: بانک‌ها و شرکت‌های کارت اعتباری از ابزارهای داده‌کاوی برای ساخت مدل‌های ریسک مالی، شناسایی تراکنش‌های متقلبانه و درخواست‌های وام و اعتبار استفاده می‌کنند. داده‌کاوی همچنین نقش کلیدی در بازاریابی و شناسایی فرصت‌های بالقوه‌ی افزایش فروش با مشتریان حال حاضر را دارد.

۵- مراقبت‌های بهداشتی در آینده: داده‌کاوی پتانسیل زیادی برای بهبود سیستم سلامت دارد؛ و با استفاده از داده‌ها و تجزیه و تحلیل آنها سعی می‌کند بهترین روش‌ها برای مراقبت‌های بهداشتی را به‌طوری که هزینه‌ها کم کاهش پیدا کنند، بیابد. محققان از روش‌های داده‌کاوی مانند پایگاه‌های داده چند بعدی، یادگیری ماشینی، محاسبات نرم، تجسم داده‌ها و آمار استفاده می‌کنند. داده‌کاوی می‌تواند برای پیش‌بینی حجم بیماران در هر رده استفاده شود. فرآیندهایی ایجاد می‌شوند که اطمینان حاصل می‌کنند بیماران مراقبت‌های مناسب را در مکان و زمان مناسب دریافت می‌کنند. داده‌کاوی همچنین می‌تواند به بیمه‌گران مراقبت‌های بهداشتی کمک کند تا ثقل و سوءاستفاده را شناسایی کرده و از آن جلوگیری کنند.

۶- پیش‌گیری از وقوع جرم: داده‌کاوی نقاط پرت و دور از نظر را به طور گسترده‌ای تحت‌نظر قرار داده و بررسی می‌کند. داده‌های جنایی شامل تمام جزئیات جرمی است که اتفاق افتاده است. داده‌کاوی الگوها و روندها را مطالعه می‌کند و رویدادهای آینده را با دقت بیشتری پیش‌بینی می‌کند.

پایگاه‌های امنیتی و پلیس‌ها می‌توانند دریابند که کدام منطقه بیشتر در معرض جرم و جنایت است، چه تعداد پرسنل پلیس باید مستقر شوند، کدام گروه سنی باید مورد هدف قرار گیرند و ...

به طور کلی می‌توان گفت که بهره‌برداری از داده‌کاوی، گستره‌ای از بخش‌های مالی برای جستجو و پردازش الگوها در بازارها، تا دولت‌ها برای شناسایی تهدیدهای امنیتی را شامل می‌شود.

این موضوع، علاوه بر کارکردهای مواجهه با مشتری مانند بازاریابی، تبلیغات، فروش و پشتیبانی مشتری شامل تولید، مدیریت زنجیره‌ی تامین، امور مالی و منابع انسانی می‌باشد. داده‌کاوی از کشف ثقل، مدیریت ریسک، برنامه‌ریزی امنیت سایبری و بسیاری دیگر از موارد استفاده‌ی حیاتی در تجارت نیز پشتیبانی می‌کند. همچنین نقش مهمی در مراقبت‌های بهداشتی، دولتی، تحقیقات علمی، ریاضیات، ورزش و... ایفا می‌کند.

در زیر چند مورد از کاربردهای آن را مشاهده می‌کنیم:

۱- رسانه‌های اجتماعی: شاید یکی از ملموس‌ترین کاربردهای داده‌کاوی در زندگی ما استفاده از شبکه‌های اجتماعی باشد. حتماً برای شما هم پیش‌آمده که بعد از مدتی گشت‌وگذار در اینستاگرام متوجه شده‌اید که تمام صفحه کاوش شما پر از موارد و موضوعاتی است که شما به آن علاقه‌مندید یا در چند وقت گذشته در مورد آن جستجو یا حتی فکر کرده‌اید!

در واقع بسترهایی مانند فیسبوک، تیک تاک، اینستاگرام، یوتیوب و توییتر مجموعه‌ای از داده‌ها را در مورد کاربران فردی جمع‌آوری می‌کنند تا در مورد ترجیحات و علاقه‌مندی‌های آنها استنباط کنند تا تبلیغاتی هدفمند برای آنها ارسال شود. همچنین از این داده‌ها برای تأثیرگذاری بر رفتار کاربران و تغییر ترجیحات آنها نیز استفاده می‌شود. خواه برای خرید و مصرف یک محصول باشد یا این که افراد در انتخابات به چه کسی رأی می‌دهند.

مزایا و محدودیت‌ها

در داده‌کاوی

۳- پیش‌بینی روندهای آینده: سیستم‌های داده‌کاوی اطلاعاتی را که از روند کار به‌دست می‌آورند را بررسی می‌کنند و می‌توانند در پیش‌بینی روندهای آینده به ما کمک کنند.

۴- کمک در تصمیم‌گیری: مردم از تکنیک‌های داده‌کاوی برای کمک به تصمیم‌گیری در بازاریابی یا کسب‌وکار استفاده می‌کنند.

محدودیت‌ها در داده‌کاوی: علی‌رغم اهمیت بسیار بالای داده‌کاوی در کسب‌وکارهای امروزی و دستاوردهای مهمی که این علم برای سازمان‌ها ایجاد می‌کند، چالش‌ها و مشکلاتی نیز در این مسیر وجود دارد. در ادامه، شماری از مهم‌ترین چالش‌های داده‌کاوی را ذکر می‌کنیم و سپس به شرح

مزایای انجام داده‌کاوی: همان‌طور که به بررسی کاربردهای داده‌کاوی در زندگی پرداختیم باید به این نتیجه رسیده‌باشیم که انجام داده‌کاوی مزایای زیادی به همراه دارد که در ادامه به چند مورد آن‌ها اشاره می‌کنیم:

۱- بازاریابی و فروش موثرتر: داده‌کاوی به بازاریابان کمک می‌کند رفتار و ترجیحات مشتری را بهتر درک کنند و در مقابل، رفتار بهتری از خود نشان دهند.

۲- خدمات بهتر به مشتریان: به لطف داده‌کاوی، شرکت‌ها می‌توانند مشکلات بالقوه‌ی خدمات مشتری را سریع‌تر شناسایی کنند و سعی در حل آن داشته باشند.



برخی از این موارد می‌پردازیم:

اصلی‌ترین چالش‌های داده‌کاوی عبارتند از:

- مسائل امنیتی و حفظ حریم خصوصی
- مواجهه با داده‌های ناقص و پراکنده
- دشواری کشف پیچیدگی‌های موجود در برخی داده‌ها
- چالش‌های روش‌شناختی
- لزوم انتخاب روش تحلیل درست برای استخراج نتایج کارآمد
- مقیاس‌پذیری الگوریتم‌ها

دشواری در ارائه مفاهیم شهودی برای برخی پدیده‌های نهفته در داده‌ها از آنجا که داده‌های خام سازمان‌ها ممکن است حاوی اطلاعات ارزشمندی از ابعاد مختلف زندگی کاربران باشند، نگهداری و حفظ امنیت این داده‌ها و خدشه وارد نشدن به حریم خصوصی کاربران طی کار بر روی داده‌ها، کار بسیار دشواری است. از سوی دیگر، داده‌های خامی که برای انجام داده‌کاوی و استخراج اطلاعات مورد نیازند، لزوماً به سادگی به دست نمی‌آیند و یا اگر بتوان به این داده‌ها دست پیدا کرد، دسته‌بندی و پاکسازی آنها از داده‌های پرت و زائد کار چندان ساده‌ای نیست. حتی ممکن است پراکندگی داده‌های به دست آمده به قدری باشد که یکپارچه و هماهنگ کردن آنها، دست‌اندرکاران داده‌کاوی را با مشکل مواجه کند.

اعتبار و کارآمدی نتایج حاصل از داده‌کاوی، وابستگی بسیاری به دقت در انتخاب روش‌ها و الگوریتم‌های درست برای تحلیل داده‌ها دارد. چنانچه ابزارها و تکنیک اتخاذ شده برای بررسی مجموعه‌ای از داده‌ها مناسب نباشد، ممکن است اعتبار نتایج حاصل از داده‌کاوی خدشه‌دار شود. از سوی دیگر، باید از الگوریتم‌هایی استفاده شود که مقیاس‌پذیر باشند تا توان پاسخگویی به حجم متفاوتی از داده‌ها در حوزه‌های مختلف سازمان را داشته باشند.

در نهایت باید گفت، گاهی ممکن است ارائه توضیحات شهودی و درک مفاهیم کشف شده در دل داده‌ها کار ساده‌ای نباشد؛ به همین دلیل، برای

استخراج درست مفاهیم از داده‌های هر حوزه، ممکن است به متخصصانی نیاز باشد که دانش زیادی در علوم داده و آن حوزه خاص مورد بررسی، داشته باشند.

همچنین اگر بخواهیم بخشی از مطالب بالا به فرض مثال مورد «مسائل امنیتی و حفظ حریم خصوصی» را با جزئیات بیشتر بررسی کنیم، می‌توان گفت:

۱) مسائل خصوصی:

داده‌کاوی اطلاعات افراد را با تکنیک‌ها و فناوری‌های مبتنی بر بازار جمع‌آوری می‌کند. این فرآیند داده‌کاوی شامل چندین عامل است اما همین عوامل، حریم خصوصی کاربر را نقض می‌کند. به همین دلیل مسائل امنیتی و امنیت کاربران کم است و در نهایت، باعث ایجاد ارتباط نادرست بین افراد می‌شود.

۲) مسائل امنیتی:

از آنجایی که داده‌های عظیمی در سیستم‌های داده‌کاوی جمع‌آوری می‌شوند، برخی از این داده‌ها که بسیار حیاتی هستند، ممکن است توسط هکرها هک شوند، همان‌طور که در مورد بسیاری از شرکت‌های بزرگ مانند فورد موتورز، سونی و غیره اتفاق افتاد.

۳) اطلاعات بی‌ربط اضافی جمع‌آوری شده:

عملکرد سیستم‌ها فضای مناسبی را برای اطلاعات مفید ایجاد می‌کند. اگرچه، یک مشکل در این مجموعه اطلاعات وجود دارد که جمع‌آوری اطلاعات می‌تواند طاقت فرسا باشد چون گاهی اطلاعاتی که مد نظر ما نیست، جمع‌آوری می‌شود.

۴) سوءاستفاده از اطلاعات:

در سیستم داده‌کاوی، امکان اقدامات امنیتی و امنیتی واقعاً حداقل است و به همین دلیل است که برخی می‌توانند از این اطلاعات برای آسیب رساندن به دیگران سوءاستفاده کنند. بنابراین، سیستم داده‌کاوی باید مسیر کار خود را تغییر دهد تا بتواند سوءاستفاده از اطلاعات را در فرآیند داده‌کاوی کاهش دهد.

کاربردهای ابزارهای تحلیلیگر داده‌کاوی و متن‌کاوی در چابک‌سازی سازمان‌های مراقبت بهداشتی و درمانی

نویسنده: مهتاب کریمی

مهم حفظ سلامت و مراقبت از حیات جامعه باید در مقابل این تغییرات با چابکی برخورد کنند. یعنی باید توانایی تحلیل شرایط و موقعیت فعلی و آتی خود را داشته و برای آینده ناشناخته‌ای که در انتظار است، تصمیم‌گیری درست و به موقعی داشته باشند. این سازمان‌ها بنا بر تنوع و تکرار فعالی‌ها از نظر داده غنی هستند، اما نیازمند تکنیک‌ها و ابزارهایی نظیر داده‌کاوی و متن‌کاوی هستند تا به کمک این ابزارها، حجم زیاد داده‌ها را تبدیل به اطلاعات مفید و دانش کنند.

دانش تولید شده به یک سازمان کمک می‌کند که در جهت ایجاد و بقا چابکی حرکت کند که در نتیجه باعث ارتقاء سلامت در سطح جامعه می‌شود.

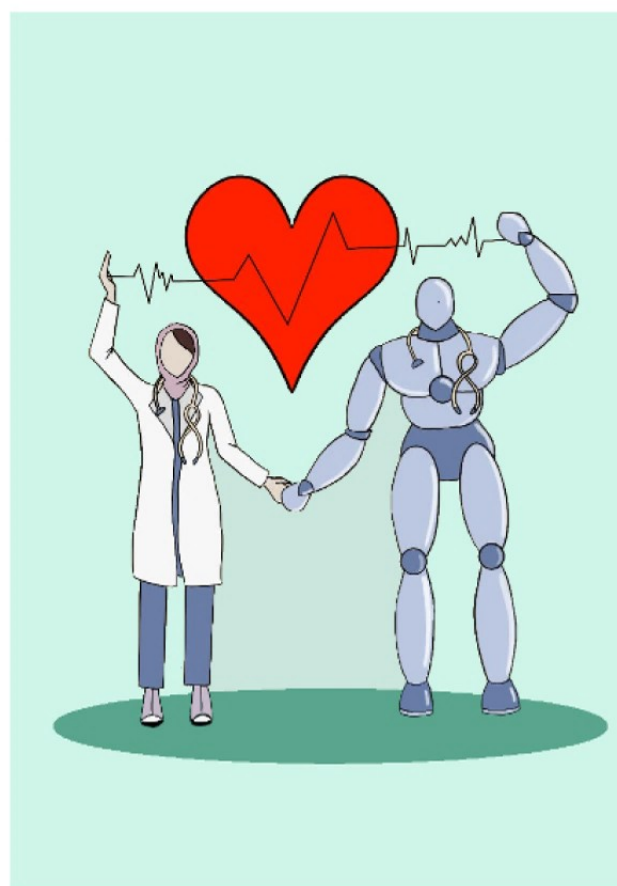
کشف یا استخراج دانش به دو صورت انجام می‌پذیرد:

۱- داده‌کاوی ۲- متن‌کاوی

۱- داده‌کاوی یا استخراج دانش از پایگاه داده‌ها فرآیند مهم شناسایی الگوهای معتبر، جدید و قابل فهم از میان انبوهی از داده‌هاست. در واقع، داده‌کاوی در مراحل مختلف و با روش‌های گوناگون اطلاعات را تبدیل به الگوها می‌کند که با تفسیر آنها به دانش می‌رسیم.

۲- متن‌کاوی به معنای استخراج خودکار اطلاعات جدید ناشناخته از منابع مکتوب گوناگون می‌باشد. تفاوت متن‌کاوی با داده‌کاوی این است که تلاش برای یافتن الگوهای جدید به جای پایگاه داده در متن‌ها صورت می‌گیرد.

به کمک این تکنیک‌ها، می‌توان به افرادی که مسئول گرفتن تصمیمات موثر در سازمان هستند، کمک کرد تا با تصمیم‌گیری‌های صحیح و به موقع باعث افزایش خلاقیت و نوآوری، دانش، خدمات‌رسانی صحیح و کاهش هزینه‌ها



تعریف واژه چابک، سرعت و قدرت در پاسخ‌گویی در هنگام مواجهه با رویدادهای داخلی و خارجی در یک سازمان است. در دنیای امروز که رقابت بسیار بالاست، چابکی نقش بسیار مهمی را ایفا می‌کند چرا که همه چیز با سرعت بالایی دستخوش تغییر است. سازمان‌های مراقبت بهداشتی و درمانی هم با توجه به برخورداری از وظیفه

مراقبت با کیفیت، همراه با کاهش هزینه‌ها می‌باشد. این سازمان‌ها می‌توانند با استفاده از داده‌کاوی و متن‌کاوی، ارتباطات بین تشخیص‌ها و آزمایشات انجام شده، تعداد مشاوره‌ها و مدت اقامت و سوءاستفاده‌ها و تقلب‌ها را شناسایی نموده و به بیشترین کارایی و کمترین مخارج در دوره‌های درمانی برسند.

در صورت تمایل به کسب اطلاعات بیشتر در این مورد، می‌توانید ویدئو همایش کاربرد هوش مصنوعی در سلامت را در [کانال آپارات نشریه صفرویک](#) تماشا کنید.

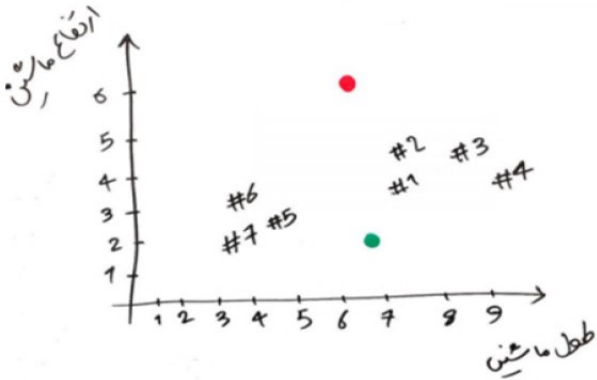
در سازمان شوند که در نهایت باعث افزایش چابکی در سازمان می‌گردد. از جمله ویژگی‌های سازمان‌های چابک، آمادگی برای پاسخ‌گویی به تغییرات و اتفاقات ناگهانی است. به کمک این ابزارها می‌توان حمله بیمای‌های محیطی، عفونی و بیوتروریسمی را در سطوح محلی، منطقه‌ای و ملی شناسایی کرد و برای پاسخ‌گویی به آن آماده بود.

از دیگر کاربردهای یک سازمان چابک، این است که به کمک آنها می‌توان عوامل خطر آفرین و موثر بر حمله و شیوع بیماری‌های مزمن را کشف و افراد در معرض خطر را شناسایی و طبقه‌بندی نمود و براساس آن، برنامه‌های مراقبتی برای درمان سریع‌تر بیماران ارائه داد.

در نهایت، می‌توان گفت هدف سازمان‌های مراقبت بهداشتی و درمانی، ارائه

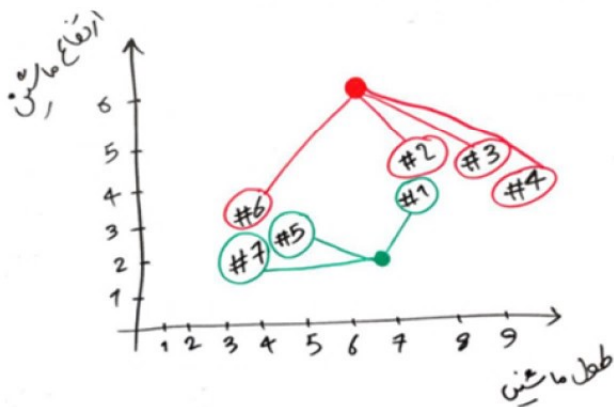
الگوریتم k-means

• سپس دو نقطه فرضی (سبز و قرمز) را در نمودار مشخص می‌کنیم. علت این که دو نقطه را انتخاب می‌کنیم این است که می‌خواهیم داده‌ها را به دو خوشه/گروه تقسیم کنیم. (تصویر ۲)



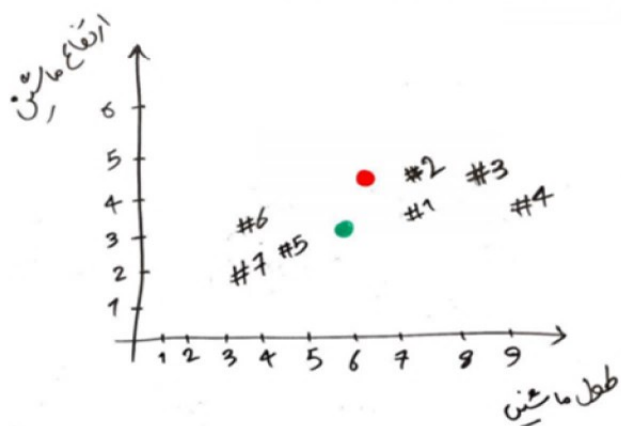
تصویر ۲

• حال فاصله تک تک داده‌ها را از نقاط سبز و قرمز محاسبه می‌کنیم. برای هر داده، اگر فاصله تا نقطه سبز کم‌تر بود، در خوشه سبز و اگر تا نقطه قرمز کمتر بود، در خوشه قرمز قرار می‌گیرد. (تصویر ۳)



تصویر ۳

• حال باید در هر خوشه میانگین داده‌ها را یافته و آن مقدار را به عنوان نقطه سبز و قرمز جدید در نظر بگیریم. (تصویر ۴)



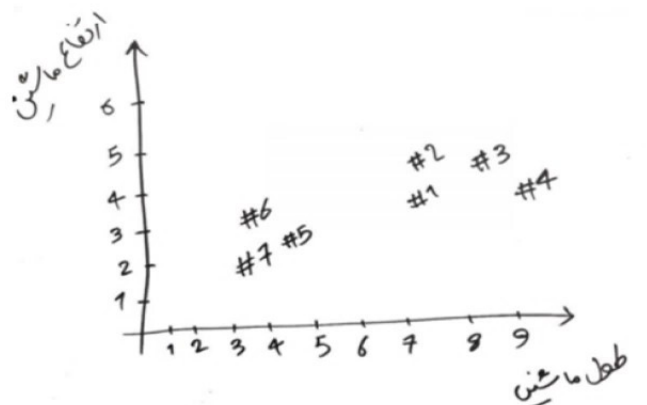
تصویر ۴

یکی از ساده‌ترین و معروف‌ترین الگوریتم‌های علم داده‌کاوی، الگوریتم یا خوشه‌بندی (k-means) است. این الگوریتم زمانی کاربرد دارد که بخواهیم برای تعدادی از داده‌ها گروه‌بندی‌های ذاتی بسازیم، بدون آن که از نوع داده‌ها یا پرچسب آن‌ها اطلاعی داشته باشیم. تعداد این گروه‌بندی‌ها (خوشه‌ها) همان مقدار k است.

برای توصیف بهتر، این الگوریتم را با یک مثال توضیح می‌دهیم:

• فرض کنید یک سری داده (طول ماشین، ارتفاع ماشین) در مورد اتومبیل‌ها در اختیار داریم. می‌خواهیم این داده‌ها را در دو گروه (k) اتوبوس‌ها و پرایدها خوشه‌بندی کنیم. برای این کار ابتدا داده‌ها را در یک نمودار $x-y$ جایگذاری و مشخص می‌کنیم. (تصویر ۱)

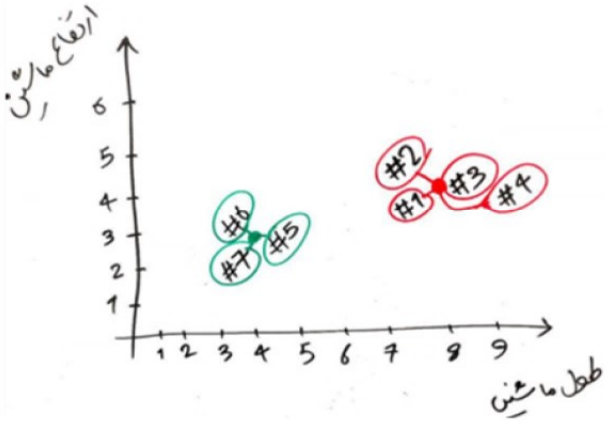
به عنوان مثال در تصویر ۱، ماشین شماره ۴ دارای طول ۹ و ارتفاع ۴ است.



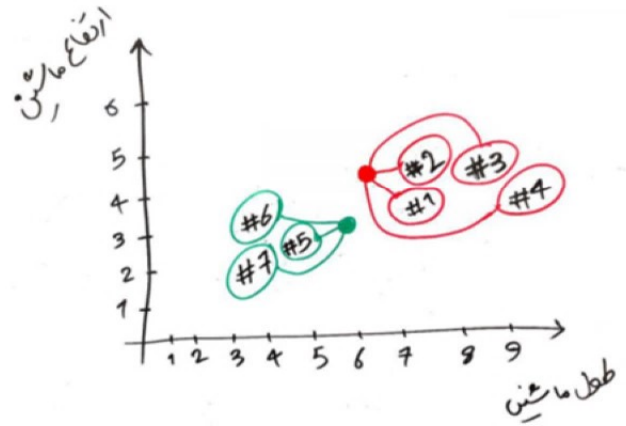
تصویر ۱

• حال همه داده‌ها در دو خوشه سبز و قرمز تقسیم کردیم بدون آن که اطلاعات مشخصی از آن‌ها داشته باشیم. (تصویر ۶)
 به همین روش الگوریتم (k-means) پیاده سازی می شود.

• دوباره فاصله تک تک داده‌ها را از نقاط سبز و قرمز یافته و گروه بندی می‌کنیم و باز هم نقاط سبز و قرمز جدید را می‌یابیم. این قدر این کار را انجام می‌دهیم که دیگر تغییری در گروه بندی‌ها به وجود نیاید. (تصویر ۵ و ۶)



تصویر ۶



تصویر ۵

از ما زنان برای مرحله بعدی مصاحبه‌های شغلی بوده است. این، وضعیت را برای زنان بدتر می‌کند زیرا آنها کم‌تر برای نقش‌های شغلی خاص مانند Data Scientist، به‌ویژه در شرکت‌های با فناوری پیشرفته که از الگوریتم‌های استخدام استفاده می‌کنند، درخواست می‌دهند. به گزارش رویترز، آمازون از الگوریتم یادگیری ماشینی به عنوان فرآیند استخدام جدید خود استفاده کرده است که به نفع زنان نبوده است. این شرکت از رزومه‌های ارائه شده قبلی از دهه قبل به عنوان داده‌های آموزشی خود برای الگوریتم استفاده کرده بود. با توجه به اینکه اکثر این رزومه‌ها از مردان تهیه شده است، برنامه دریافت که کاندیداهای مرد ارجحیت دارند و زنان را کنار می‌گذارد. این بیش‌تر به این موضوع مربوط می‌شود که چگونه می‌توان از داده‌های جانبدارانه برای اتخاذ تصمیمات نادرست استفاده کرد و نابرابری را در نیروی کار اعمال کرد.

برای این‌که زنان کنار گذاشته نشوند، تعداد زنان در صنعت STEM بسیار مهم است. با افزایش تعداد زنان در این صنایع، اطمینان حاصل می‌شود که زنان به دلیل الگوریتم‌های از پیش تعیین شده کنار نمی‌روند و به نسل جوان فرصت بهتر و امید بیش‌تری برای پیشرفت در مشاغل STEM می‌دهد. STEM مخفف Science و Technology و Engineering و Mathematics است. (علوم تجربی، تکنولوژی، مهندسی و ریاضیات)

وقتی صحبت از علم داده به میان می‌آید، بسیاری از افراد این مسیر شغلی را به‌عنوان صنعتی برای مردان می‌دانند و این دیدگاه زنان را بیش‌تر و بیش‌تر از این حرفه دور می‌کند.

مشخص است که اکثر افرادی که جمع‌آوری، تجزیه و تحلیل، تولید خروجی و تصمیم‌گیری می‌کنند، مرد هستند. طبق گزارش TechRepublic، در سال ۲۰۱۹ زنان ۱۸ درصد از مشاغل علم داده در ایالات متحده را تشکیل می‌دهند. این ارقام در کشورهای کم‌درآمد که در آن زنان کم‌تر به تحصیلات و مدارک بهتر در علوم فناوری، مهندسی و ریاضیات (STEM) دسترسی دارند، حتی بیش‌تر کاهش می‌یابد.

بر اساس گزارش بانک جهانی، در سال ۲۰۱۹، ۸۳/۰۲ درصد از زنان باسواد بودند در حالی که ۸۹/۹۳ درصد مردان باسواد بودند. در نتیجه، زنان اغلب برای تعامل موثر با فناوری‌های دیجیتال تلاش می‌کنند. ما از قبل می‌دانیم که عدم تعادل تعداد کارکنان از نظر جنسیت در مشاغل STEM وجود دارد و این زنان را از ورود به این صنعت ناامید می‌کند. سطح آموزش ارائه شده به زنان در مشاغلی مانند Data Scientist نیز کم‌تر است، و موفقیت زنان را در ایجاد مشاغل و کسب حرفه‌های پردرآمد دشوارتر می‌کند و تأکید بیشتری بر شکاف دستمزد بر اساس جنسیت دارد.

همه ما می‌دانیم که الگوریتم‌های استخدام احتمالاً دلیل انتخاب نشدن برخی

درس و ابزارها

در داده‌کاوی

در زمینه‌های مختلف مورد نیاز است تا الگوهای لازم را بتوان از این داده‌ها استخراج کرد. عمده تفاوت ابزارهای داده‌کاوی در حجم داده‌ها، نوع آنالیز و استخراج داده‌ها و روش کار است.

برخی از نرم‌افزارهای این حوزه جزو ابزارهای برتر داده‌کاوی شناخته شده‌اند که هر سال بنا بر تغییرات فرآیند داده‌کاوی به روز می‌شوند. این ابزارها عموماً بصورت رایگان قابل استفاده و منبع باز هستند. در سال ۲۰۲۰ تعداد بیشتری نرم‌افزار به عنوان ابزار داده‌کاوی برتر شناخته شد.

با توجه به اهمیت داده‌کاوی در دنیای کسب‌وکارهای بزرگ و حرفه‌ای، ابزارها و نرم‌افزارهای متعددی برای این منظور طراحی و توسعه داده شده‌اند. برخی از این ابزارها، رایگان و متن باز هستند و برخی دیگر به صورت بسته‌های نرم‌افزاری تجاری ارائه شده‌اند. از این میان، تعدادی از مواردی که بیشتر مورد استفاده هستند، در ادامه آمده است.

بدون شک، زبان برنامه‌نویسی و بسته نرم‌افزاری پایتون و R از مهم‌ترین و کارآمدترین ابزارها در زمینه تحلیل و استنتاج آماری و انجام انواع محاسبات است.



درس الزامی: داده‌کاوی محاسباتی، الگوریتم‌های پیشرفته، داده‌کاوی درس انتخابی: ریاضیات یادگیری، بهینه‌سازی محدب، بهینه‌سازی ترکیباتی، یادگیری ماشین، یادگیری ماشین آماری، داده‌کاوی پیشرفته، متن‌کاوی و وب‌کاوی، انتخاب ویژگی و استخراج ویژگی، گراف‌کاوی، مدل‌های گرافی احتمالاتی، شبکه‌های پیچیده، دیدار سازی داده‌ها، شناسایی دور افتاده‌ها، مدل‌سازی و پردازش مه داده‌ها، یادگیری ژرف، مباحث ویژه در داده‌کاوی.

یکی از مهم‌ترین درس‌ها در داده‌کاوی درس مربوط به آمار و احتمال است که در این رشته کاربردی می‌باشد. همچنین درس داده‌کاوی پیش نیاز سایر درس‌ها می‌باشد پس از این نظر اهمیت ویژه‌ای دارد.

ابزارها: وظایف ابزارهای داده‌کاوی چیست؟

ابزارهای داده‌کاوی مجموعه‌ای از تکنیک‌هایی هستند که الگوریتم‌های خاص، آنالیز آماری، هوش مصنوعی و سیستم‌های پایگاه داده را به کار می‌برند تا داده‌ها را از ابعاد و دیدگاه‌های مختلف تجزیه و تحلیل کنند. هدف این نرم‌افزارها کشف الگوها، روندها و گروه‌بندی‌های خاص در بین مجموعه‌های بزرگ داده و تبدیل داده‌ها به اطلاعات تصفیه شده‌تر است.

از ابتدای روند داده‌کاوی ابزارهای مختلفی برای انجام این کار به وجود آمد که روز به روز به‌روزرسانی می‌شوند. ابزارهای جدیدتر علاوه بر وظیفه تجزیه و تحلیل، مدیریت داده‌ها را نیز برعهده دارند. این ابزارها از مدل‌های مختلفی مانند مدل پردازش داده و تصویر سازی با استفاده از پایگاه‌های داده، فرآیند داده‌کاوی را به خوبی انجام می‌دهند. ابزارهای داده‌کاوی به کمک الگوها و روابط مشخصی حجم بالایی از اطلاعات را از پایگاه‌های داده استخراج و تجزیه تحلیل می‌کنند.

به دلیل حجم بالای داده‌های موجود تصمیم‌گیری کاربردی و هوشمندانه‌تری



R Studio®

استفاده می‌شود.

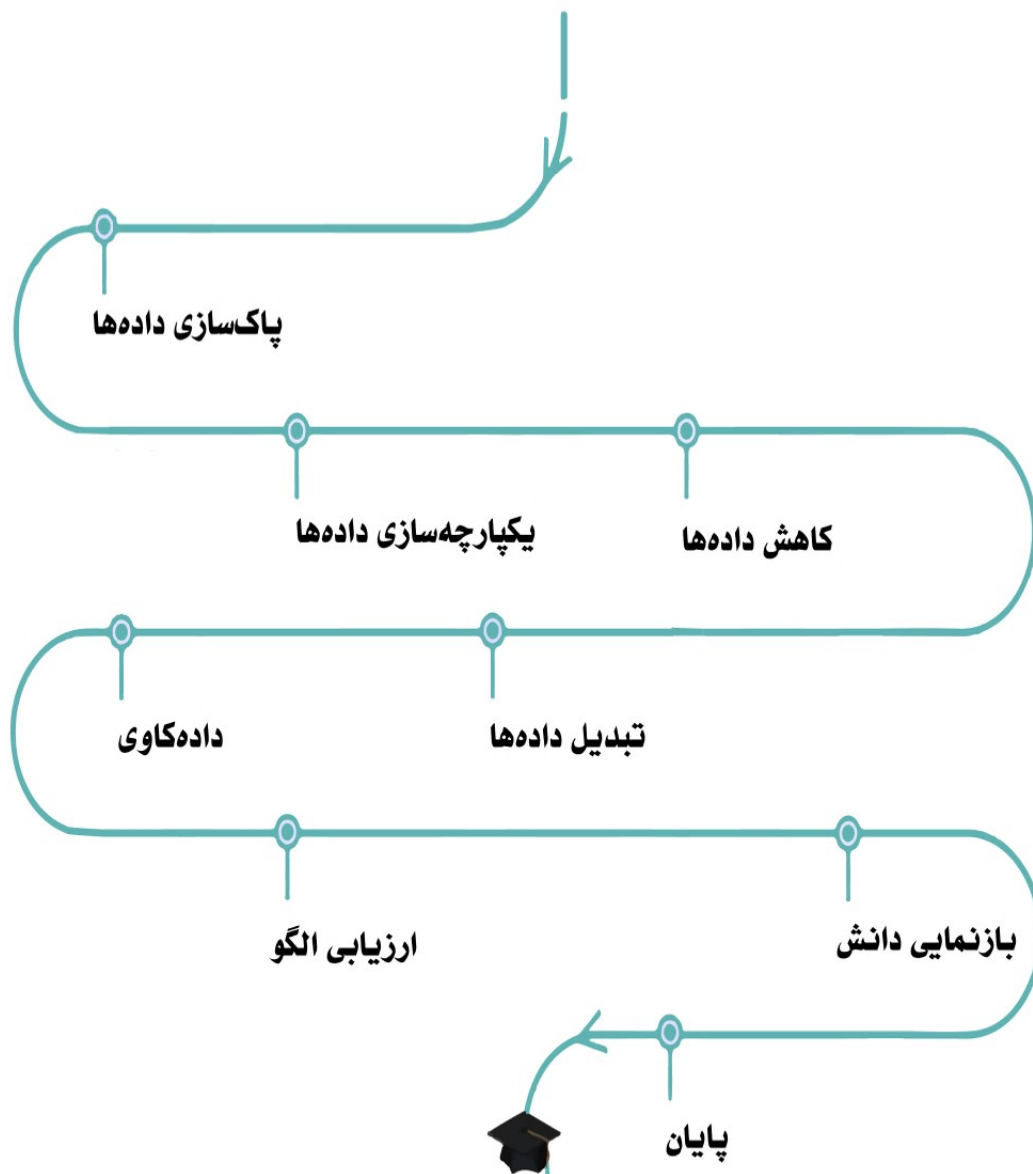
می‌توان از این زبان برنامه‌نویسی برای پیش‌بینی نتایج، خودکارسازی وظایف، ساده‌سازی فرآیندها و ارائه بینش‌های هوش تجاری استفاده کرد. امکان کار با داده‌ها در پایتون وجود دارد، اما تعداد زیادی کتابخانه متن باز وجود دارد که کار داده‌کاوی با پایتون را خیلی ساده‌تر می‌کند. مطمئناً نام برخی از این کتابخانه‌ها را شنیدماید در این بخش یکی از مهم‌ترین کتابخانه‌های پایتون برای داده کاوی ارائه شده است که شامل زمینه‌هایی مانند پردازش داده‌ها، مدل سازی و تجسم داده‌ها است.

کتابخانه NumPy: این کتابخانه یک ابزار عالی برای محاسبات علمی و انجام عملیات اولیه و پیشرفته مبتنی بر آرایه است. این کتابخانه بسیاری از ویژگی‌های مفید را برای انجام عملیات روی آرایه‌ها و ماتریس‌های پایتون را ارائه می‌دهد. همچنین به پردازش آرایه‌هایی که مقادیر یک نوع داده را ذخیره می‌کند، کمک می‌کند و انجام عملیات ریاضی روی آرایه‌ها و بردار سازی آن‌ها را آسان‌تر می‌کند. در حقیقت، برداری از عملیات ریاضی بر روی نوع آرایه NumPy باعث افزایش عملکرد برنامه‌نویسی می‌شود و زمان اجرا را تسریع می‌کند.

نرم افزار RStudio برنامه‌ای است که برای انجام محاسبات آماری از ساده‌ترین محاسبات تا پیشرفته‌ترین آن‌ها مورد استفاده قرار می‌گیرد و از طرفی هم آمار بخش جدایی ناپذیری از علم داده و فرآیند داده‌کاوی محسوب می‌شود. پس این برنامه بدون شک می‌تواند یکی از بهترین ابزارها برای انجام امور مربوط به داده‌کاوی باشد. علاوه بر این چون R خود یک زبان برنامه‌نویسی هم محسوب می‌شود، با انجام کدنویسی و ساخت فرآیندهای تکرارپذیر، به راحتی می‌توان سرعت کارهای مربوط به داده‌کاوی را افزایش داد و این مورد هم خود دلیل دیگری برای محبوبیت این زبان و برنامه در بین متخصصین علم داده محسوب می‌شود.

همچنین زبان برنامه‌نویسی پایتون، یکی از بهترین زبان‌های برنامه‌نویسی داده کاوی است. پایتون به دلیل سادگی، انعطاف بالا، همه منظوره بودن و ایجاد برنامه‌های کاربردی و تحلیل داده مورد توجه همگان قرار گرفته است. این زبان محبوب مورد استفاده اکثر برنامه‌نویسان بوده و با داشتن کتابخانه‌های مختلف باعث شده که برنامه‌نویسان زیادی به این زبان روی بیاورند. پایتون یکی از محبوب‌ترین زبان‌هایی است که توسط دانشمندان داده و توسعه‌دهندگان نرم‌افزار به طور یکسان برای کارهای داده‌کاوی و علم داده

مراحل کاوش داده و مهارت‌های ضروری



برای استفاده درست از الگوهایی که یک داده‌کاو در داده‌های جمع‌آوری شده پیدا می‌کند، باید حس تجاری دقیقی داشته باشد. تجزیه و تحلیل داده‌ها فقط با دیدگاه روشن نسبت به اهداف و آینده‌نگری امکان‌پذیر است. همچنین متخصص داده‌کاوی باید بتواند یافته‌های تخصصی را طوری ارائه کند که مخاطبان بتوانند آنها را درک کنند. پس باید مهارت‌های قوی در صحبت کردن در عموم و توانایی انتقال نتایج به مخاطب را داشته باشد.

یک متخصص داده‌کاوی به ترکیبی منحصر به فرد از مهارت‌های فنی، تجاری و فردی نیاز دارد. از جمله مهارت‌های فنی که یک داده‌کاو باید به آنها تسلط داشته باشد، می‌توان به آشنایی با ابزارهای تجزیه و تحلیل داده، توانایی استفاده از زبان‌های برنامه‌نویسی جاوا، پایتون و پرل و تجربه کار با انواع سیستم‌عامل‌ها مخصوصاً لینوکس اشاره کرد.

معرفی کتاب و دوره

در داده‌کاوی

کتاب:

Data Mining: Concepts and Techniques Book by Jiawei Han

این کتاب یکی از بهترین کتاب‌های جامع در داده‌کاوی است که در ۱۳ فصل گردآوری شده است. این کتاب به طوری نوشته شده است که هم برای افراد مبتدی و هم افراد حرفه‌ای مناسب و قابل استفاده می‌باشد.

یکی دیگر از کتاب‌های مفید در حوزه داده‌کاوی کتاب «تبدیل شدن به یک مدیر داده» می‌باشد که نوشته الکس جی گاتمن و جردن گلد میر است.

دوره تخصصی داده‌کاوی کورسرا:

- به زبان انگلیسی، به صورت برخط، سطح متوسط.
- مر بیان:

جان سی هارت، چنگ شیانگ ژای، جیاوئی هان

- مباحث و سرفصل‌ها:

تجسم داده‌ها، بازیابی متن و موتور جستجو، متن کاوی و تجزیه و تحلیل، کشف الگو در داده‌کاوی، تجزیه و تحلیل خوشه‌ای در داده‌کاوی، پروژه داده‌کاوی

- مهارت‌هایی که به دست خواهید آورد:

الگوریتم‌های خوشه‌بندی داده‌ها، استخراج متن، تجسم داده‌ها (DataViz)، داده‌کاوی، نرم‌افزار تجسم داده‌ها، نرم افزار تابلو، مجازی‌سازی داده‌ها، بازیابی اطلاعات (IR)، بازیابی سند، فراگیری

ماشین، سیستم توصیه‌گر، مدل‌های احتمالی

- در پایان دوره گواهینامه داده می‌شود.

- لینک مستقیم دوره: [کلیک کنید](#).

دوره:

داده‌کاوی Microsoft Data Mining دانشگاه صنعتی شریف:

- به زبان فارسی، این دوره شامل ۵ بخش می‌باشد
- نوع دوره: به صورت حضوری و برخط
- مدت زمان دوره: ۴۰ ساعت
- پیش نیاز: ندارد اما بلد بودن پایتون به یادگیری شما کمک می‌کند.
- مباحث و سرفصل‌ها:

نصب و پیاده‌سازی محیط‌های عملیاتی، مفاهیم پایه داده‌ها و ریاضی و آماری، نمایش داده‌ها، طبقه‌بندی و رگرسیون و الگوریتم‌های مختلف آن، خوشه‌بندی و الگوریتم‌های مختلف آن.

- مهارت‌هایی که به دست خواهید آورد:

نصب و ایجاد محیط اولیه در ادیتورهای مختلف، آشنایی و کار با کتابخانه‌های مختلف، معرفی روش‌ها و مراجع جمع‌آوری داده‌ها و استفاده از آن مثل پیش‌بینی خرید کاربر با توجه به رفتار او در فروشگاه اینترنتی بررسی درخت‌های تصمیم، کاربرد و آشنایی با روش‌های عملی خوشه‌بندی، آشنایی با نمونه داده‌های طبقه‌بندی و کاربردهای آن و ...

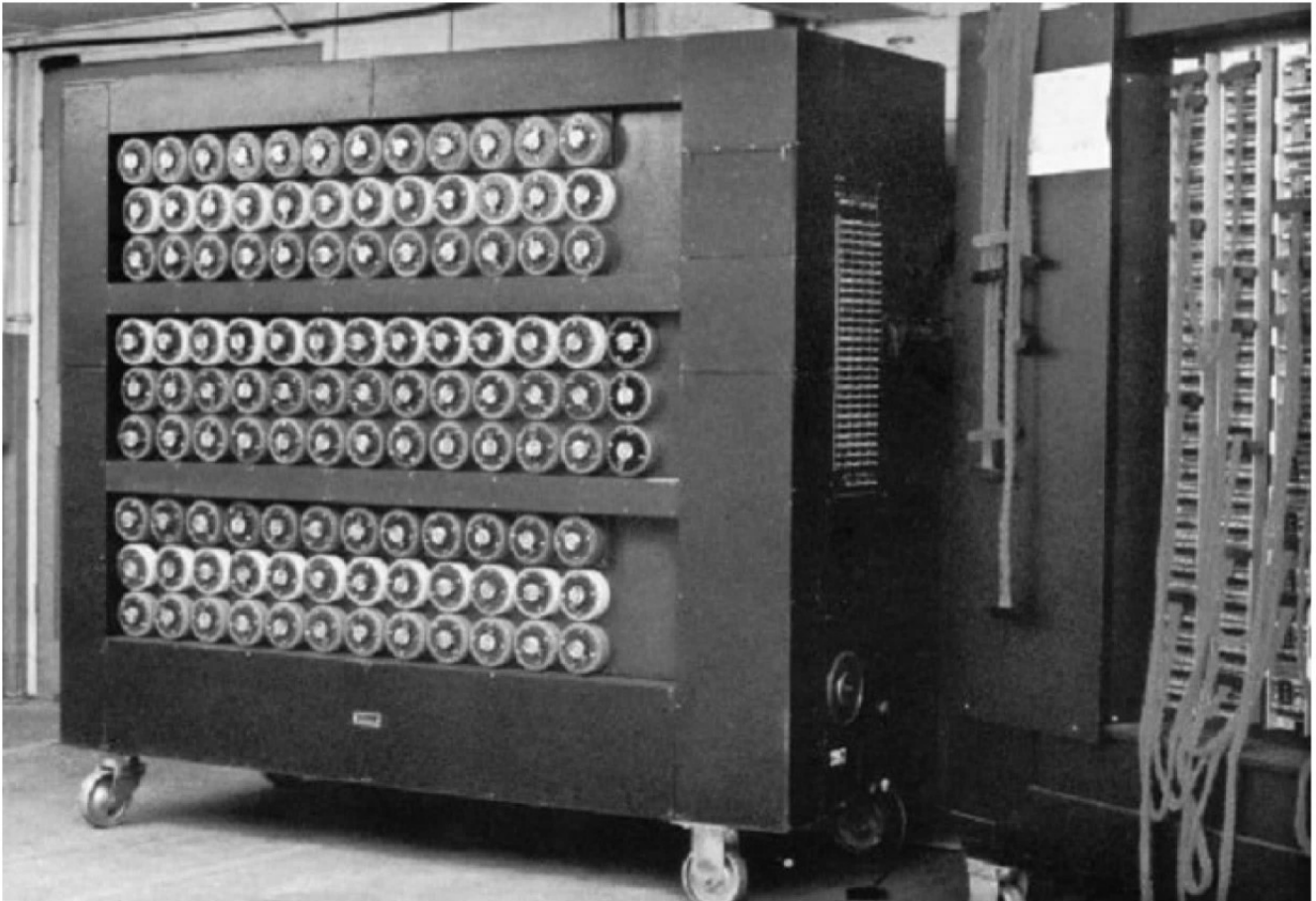
- در پایان دوره گواهینامه داده می‌شود.

- لینک مستقیم دوره: [کلیک کنید](#).

مجموعه کتب مکتب ائمه



0011



هوش مصنوعی != هوش طبیعی

اولین جرقه ایجاد هوش مصنوعی سوالی بود که یکی از فیلسوفان خارق‌العاده به نام «آلن تورینگ» بیان کرد. او می‌گوید انسان‌ها از طریق اطلاعات موجود به عنوان دلیلی استفاده می‌کنند تا به حل مشکلات و تصمیم‌گیری بپردازن؛ پس چرا ماشین‌ها نتوانند؟

فلسفه وجود هوش مصنوعی شاید به زمان جنگ جهانی دوم زمانی که نیروهای آلمانی ماشینی برای رمزگذاری و ارسال پیام‌های امن به نام enigma طراحی نمودند، بازگردد؛ انگلیسی‌ها در رمزنگاری این پیام‌ها دچار مشکل می‌شدند در اینجا بود که آلن تورینگ ماشینی به اسم bombe طراحی کرد تا به سادگی کدها را رمزگشایی کند. در واقع اینجا بود که آلن تورینگ به اهمیت هوش مصنوعی پی برد.

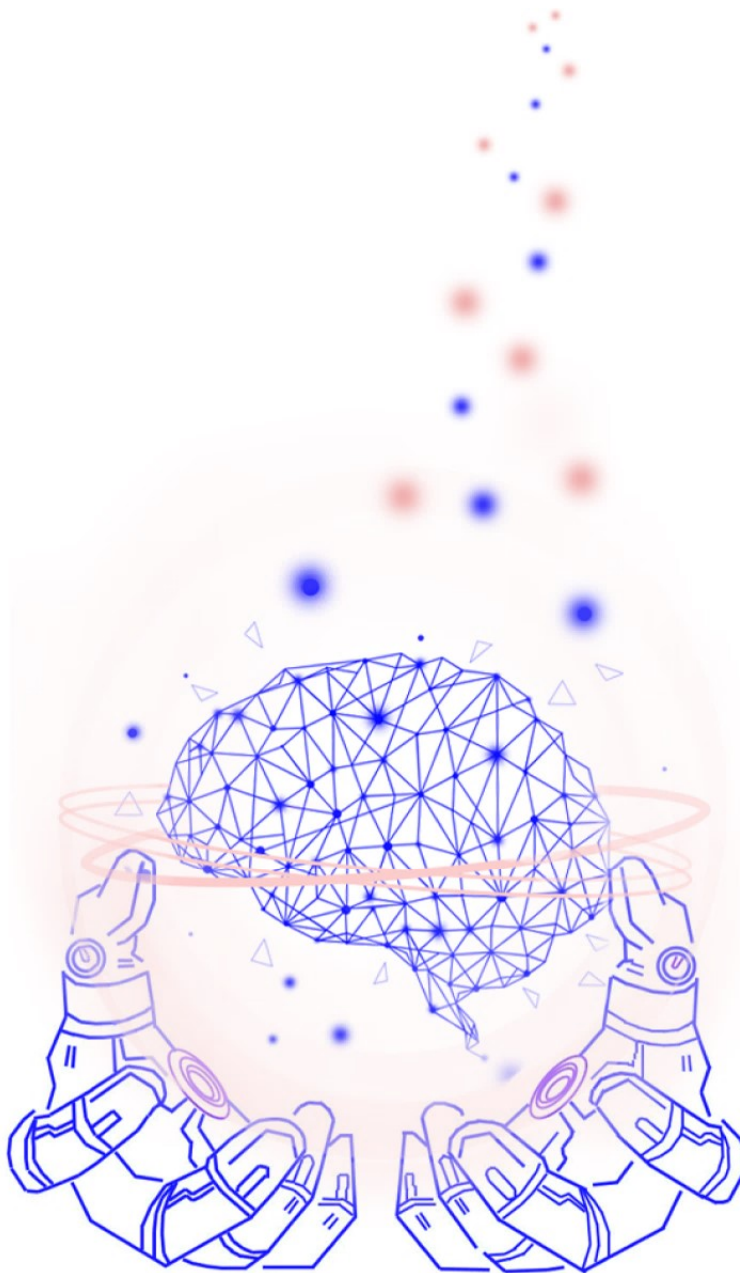
هوش مصنوعی اشاره به نرم‌افزارهایی دارد که علاوه بر تجزیه و تحلیل داده‌ها، می‌تواند قدرت تصمیم‌گیری داشته باشد. در واقع تقلیدی از هوش انسان متخصص است. سیستم‌هایی که از هوش مصنوعی استفاده می‌کنند، به‌طور آگاهانه بخش یادگیری ماشینی را فعال کرده‌اند. نتیجه عملکرد این سیستم‌های مبتنی بر هوش مصنوعی، تجزیه و تحلیل الگوریتم‌ها و ارائه نتایج قطعی و معتبر است. همواره این پتانسیل و مزیت برای هوش مصنوعی وجود دارد که به‌طور خودکار نشانه‌هایی را برای تعیین تصمیم نهایی، درک و تحلیل کند. در مواردی که سیستم دچار خطا می‌شود این امکان وجود

در نیمه اول قرن بیستم داستان‌های علمی تخیلی آن روزها، شروع اولین تفکرها درباره هوش مصنوعی در ذهن کودک تا بزرگسال و دانشمند را رقم زد.

در دهه ۱۹۵۰ میلادی، «شاه دنسلی» از دانشمندان، ریاضی‌دانان و فیلسوفانی بود که دارای ذهنی مشتاق و مجذوب برای پیاده کردن افکار خود درباره هوش مصنوعی داشت.

تقریباً بعد از گذشت یک دهه نمادهای هوش مصنوعی، تبدیل به رشته‌ای شد که اکنون می‌شناسیم. آلن تورینگ یکی از ریاضی‌دانانی بود که آزمونی را پیشنهاد داد که توانایی ماشین را برای تکثیر اعمال انسان، به میزانی که قابل تمایز از رفتار خود انسان نباشد، اندازه‌گیری می‌کرد. کمی پس از این دهه، رشته هوش مصنوعی در طول کنفرانس تابستان در دانشکده (دارتموث) در میانه‌ی دهه ۱۹۵۰ تأسیس شد و «جان مک کارتی» دانشمند علوم شناختی و کامپیوتر، در این کنفرانس عبارت «هوش مصنوعی» را برای اولین بار به کار برد و «آلن نیوول» و «هربرت سیمون» از دانشگاه کارنگی ملون (CMU)، «جان مک‌کارتی» و «ماروین مینسکی» از موسسه فناوری ماساچوست (MIT) و «آرتور ساموئل» از شرکت (IBM)، به بنیان‌گذاران و رهبران پژوهش‌های هوش مصنوعی جهان مبدل شدند.

دارد که توسط متخصصین تصحیح گردد و این می‌تواند یک الگوریتم جدید برای جلوگیری از تکرار شدن خطا باشد. ارزش هوش مصنوعی بر پایه همین یادگیری، ارتقاء داده‌ها و بالا رفتن قدرت تجزیه و تحلیل است. از دیدگاه تاریخی نام هوش مصنوعی در سال (۱۹۶۵) میلادی به عنوان یک دانش جدید ابداع گردید. البته فعالیت در این زمینه از سال (۱۹۶۰) میلادی آغاز شده بود. بیشتر کارهای پژوهشی اولیه در هوش مصنوعی بر روی انجام ماشینی بازی‌ها و نیز اثبات قضیه‌های ریاضی با کمک رایانه‌ها بود. در آغاز چنین به نظر می‌آمد که رایانه‌ها قادر خواهند بود چنین فعالیت‌هایی را تنها با بهره گرفتن از تعداد بسیار زیادی کشف و جستجو برای مسیرهای حل مسئله و سپس انتخاب بهترین روش برای حل آن‌ها به انجام رسانند.



زیرشاخه‌ها و کاربردهای

هوش مصنوعی

زیرشاخه‌های هوش مصنوعی:

(۱) یادگیری ماشین:

یادگیری ماشین یکی از شاخه‌های پرطرفدار هوش مصنوعی است. یادگیری ماشین این امکان را به سیستم‌ها می‌دهد تا به صورت خودکار و بدون نیاز به برنامه‌نویسی صریح آموزش ببینند. تمرکز متخصصین این حوزه معطوف به آن است که ماشین‌ها داده‌های ورودی را ذخیره کنند و سپس برای یادگیری خود ماشین، از آن داده‌ها استفاده کنند.

از کاربردهای یادگیری ماشین می‌توان به تشخیص چهره، تشخیص گفتار، سامانه توصیف‌گر و خدمات مالی اشاره کرد.

(۲) بینایی ماشین:

حوزه بینایی ماشین از گسترده‌ترین حوزه‌های هوش مصنوعی است. به کمک این علم، ماشین قادر به تشخیص اشیاء می‌شود. سیستم‌های بینایی ماشین به گونه‌ای برنامه‌ریزی شده‌اند که وظایفی همچون شمارش اشیاء، خواندن شماره سربال و ... را انجام دهد.

(۳) پردازش زبان طبیعی (NLP):

پردازش زبان طبیعی، سیستم‌های کامپیوتری را قادر می‌سازد تا زبان انسانی مانند فارسی، انگلیسی و ... را درک و پردازش کند. پردازش زبان طبیعی این امکان را می‌دهد تا انسان و کامپیوتر بتوانند با یکدیگر ارتباط برقرار کنند. یکی از اهداف مهم متخصصان حوزه هوش مصنوعی آن است که نرم‌افزارهایی را طراحی کنند که قدرت درک و فهمیدن زبان انسانی را داشته باشد. برای مثال می‌توان به Siri و یا Google اشاره کرد که که کاربران به

راحتی می‌توانند سوالات خود را به صورت صوتی از این دو سیستم بپرسند.

(۴) شبکه عصبی:

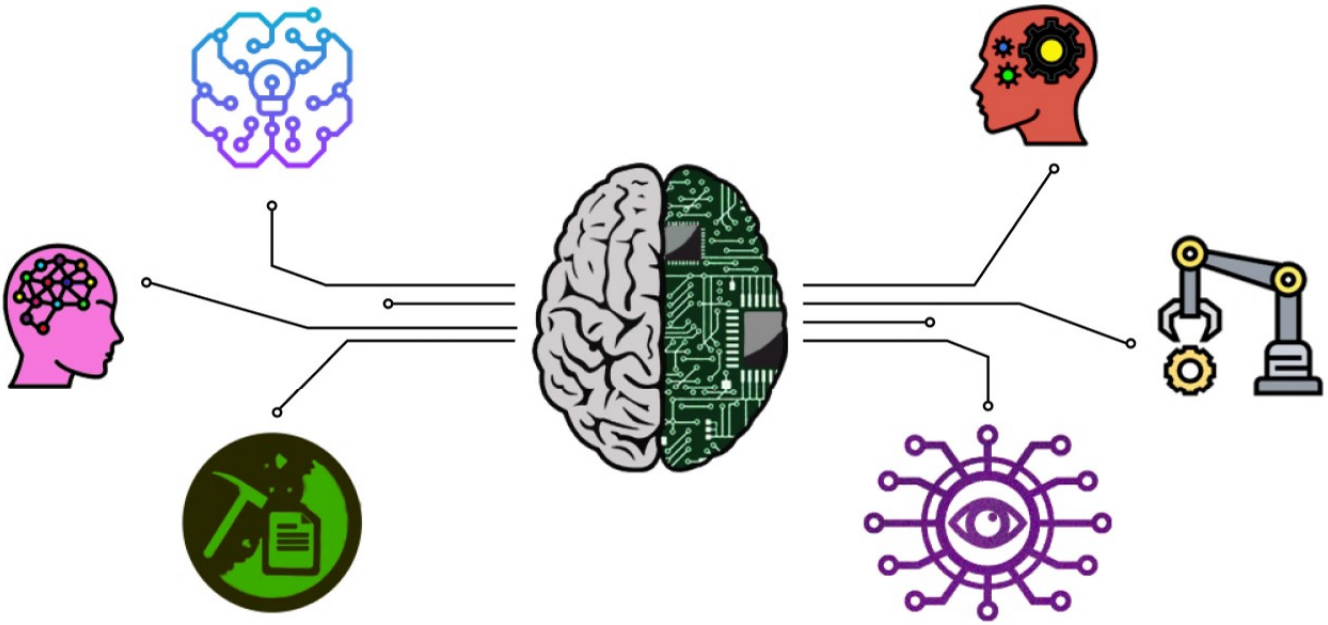
شبکه عصبی یکی از درونی‌ترین لایه‌های هوش مصنوعی است. با استفاده از الگوریتم‌های شبکه عصبی می‌توان مدل‌های پیچیده و مختلف را بررسی کرد و با شبیه‌سازی مغز انسان و حیوان می‌توان به الگوهای جدید و شناخته نشده‌ای دست یافت. از کاربرد آن می‌توان به تطابق چهره، تشخیص دست‌خط و راندن خودکار اتومبیل اشاره کرد.

(۵) سیستم‌های خبره:

سیستم‌های خبره، یک نرم‌افزار مبتنی بر هوش مصنوعی است. برخلاف سایر شاخه‌ها که بر روی داده‌ها کار می‌کنند، سیستم‌های خبره، دانش انسانی را یاد می‌گیرند و این دانش را ذخیره و از آن در تصمیم‌گیری‌ها استفاده می‌کنند. این سیستم‌ها به جای استفاده از برنامه‌نویسی‌های رایج، منطق if-then را برای حل مسائل پیچیده به کار می‌برند. از کاربردهای سیستم‌های خبره می‌توان به بررسی وام‌های بانکی، پردازش سیستم‌های پزشکی، کشف و شناسایی ویروس اشاره کرد.

(۶) ریباتیک:

ریباتیک یکی از شاخه‌های جالب هوش مصنوعی است. هدف آن طراحی، ساخت و استفاده از ربات‌ها می‌باشد. این رشته از ادغام سه رشته مهندسی مکانیک، مهندسی برق و علوم کامپیوتر به وجود آمده است. ربات‌ها می‌توانند به انسان‌ها در انجام کارهای تکراری و یا پرخطر کمک کنند. برای مثال می‌توان از آن‌ها در خطوط تولید کارخانه‌ها، صنعت خودروسازی، تجهیزات پزشکی و ... استفاده کرد.



- هوش مصنوعی در تصفیه آب:

یکی از بزرگ‌ترین دغدغه‌های مردم جهان در شرایط فعلی تصفیه آب می‌باشد. بر اساس آخرین گزارشات سازمان یونسف درباره تأثیر ویروس کرونا بر پاکیزگی آب آشامیدنی و تصفیه آب، نشان می‌دهد که دسترسی به آب آشامیدنی و تمیز به صورت چشمگیری کاهش پیدا کرده است.

شرکت‌های آب و فاضلاب با کمک هوش مصنوعی می‌توانند از گزارشات لحظه‌ای میزان آلودگی آب استفاده کنند. این داده‌ها به کارمندان این شرکت این امکان را می‌دهد تا آلاینده‌های سمی موجود در آب را شناسایی کنند و باعث بهبود فرایند تصفیه بشوند. همچنین می‌توانند از اثرات مخرب بر محیط زیست نیز جلوگیری کنند.

اگر بخواهیم علاوه بر مطالب فوق که تقسیم بندی زیرشاخه‌های این رشته بود مثالی از کاربرد عملی این رشته در زندگی واقعی بیاوریم، می‌توان به مطالب زیر اشاره کرد:

- هوش مصنوعی در کشاورزی:

به لطف هوش مصنوعی صنعت کشاورزی وارد مرحله جدیدی از تکامل شده است. برای مثال پهپادهای مجهز به هوش مصنوعی امکان سم‌پاشی خودکار سموم یا کود را به طور یکنواخت در سراسر یک مزرعه فراهم می‌کند. تشخیص اشیاء می‌تواند کار بسیار خوبی در شناسایی علف‌های هرز و تمایز آن‌ها از محصولات کشاورزی انجام دهد.

چالش‌های حل شده و حل نشده در هوش مصنوعی

علاوه بر کاربردهای گسترده این رشته هنوز هم مشکلات و محدودیت‌هایی در رابطه با برخی مسائل وجود دارد که برخی از این مشکلات حل و برخی هنوز حل نشده باقی مانده‌اند در زیر به برخی از آنها اشاره می‌کنیم:
چالش حل شده:

تا چندین سال پیش، جستجوی یک کالا بدون دانستن نام آن یک چالش بزرگ بود. اما در حال حاضر هوش مصنوعی به این چالش پاسخ داده و به ما این امکان داده شده تا بتوانیم از طریق عدسی دوربین (لنز) تلفن همراه خود، درباره اشیاء مورد نظر اطلاعات کسب کنیم. شرکت گوگل در سال ۲۰۱۷ سرویس جدیدی به نام Google Lens را معرفی کرد؛ گوگل در معرفی این سرویس می‌گوید: Search what you see؛ هر آنچه را می‌بینید، جستجو کنید. این سرویس به عنوان یک موتور جستجو برای دنیای واقعی شناخته می‌شود. اساس کار این سرویس، هوش مصنوعی و یادگیری ماشین است. ابتدا گوگل تصویر شما را روی سرویس‌دهنده‌ها (سرورها) آپلود می‌کند؛ در ادامه با شبکه‌های عصبی مصنوعی، تصویر برای تشخیص محتویات تحلیل می‌شود. بعد از درک محتویات، اطلاعات مرتبط و مفید در آن زمینه به کاربر ارائه می‌شود. الگوریتم Google Lens به این صورت است که اشیاء موجود در تصویر شما را با تصاویر دیگر مقایسه می‌کند و آن تصاویر را بر اساس شباهت و ارتباط آنها با اشیاء در تصویر

اصلی، رتبه‌بندی می‌کند. این سرویس همچنین از درک خود از اشیاء در تصویر شما برای یافتن سایر نتایج مرتبط از اینترنت استفاده می‌کند. شما می‌توانید از google lens در سفرهایتان، انتخاب بهترین غذا در منو، کسب اطلاعات درباره گونه‌های مختلف گیاهان یا حیوانات، خرید پوشاک، تبدیل عکس به متن و سایر نیازهای روزمره‌تان استفاده کنید.

همچنین یکی از چالش‌های حل نشده در هوش مصنوعی به شرح زیر است:

گروهی از پژوهشگران حین انجام تحقیقاتی در زمینه یادگیری ماشین، با سؤالاتی مواجه شده‌اند که ارتباط تنگاتنگی با مسئله‌ای حل‌نشده در ریاضیات دارد. این مسئله به «فرضیه‌ی پیوستار» معروف است. در دهه ۱۹۳۰ «کورت گودل»، ریاضی‌دان اتریشی، اولین بار ادعا کرد این مسئله حل‌نشده است.

مسئله‌ای که این پژوهشگران با آن رویه‌رو بودند، مسئله «یادگیری» نام دارد. این مسئله بررسی می‌کند آیا می‌توان با استفاده از داده‌های محدود، الگوریتمی برای حدس‌زدن الگوها یافت یا خیر؟ طبق مقاله‌ای که ۷ ژانویه (برابر با ۱۷ دی) در مجله (Nature Machine Intelligence) منتشر شد، این مسئله صورت جدیدی از فرضیه‌ی اثبات‌نشده پیوستار در ریاضیات است.

هوش مصنوعی

ژنتیک

نویسندگان: دکتر داریوش فرهود، هانیه پورکلهر

ب) هوش مصنوعی قادر به تسریع فرآیند تعیین توالی ژنتیکی و ویرایش آن خواهد بود. ابزار جدید گوگل با نام Variant Deep. از جدیدترین تکنیک‌های هوش مصنوعی برای تبدیل تصویری به HTS (Human Terrain System) دقیق‌تر از یک ژنوم کامل بهره می‌برد. از زمان ظهور HTS در اواسط دهه ۲۰۰۰، این ابزار گوگل قادر به تشخیص دادن جهش‌های ژنتیکی کوچک از میان خطاهای تصادفی بود.

با این‌که امروزه می‌توانیم توالی ژن‌ها را به‌سرعت بازخوانی کنیم؛ اما هنوز در مورد این‌که این ژن‌ها چه اطلاعاتی را در اختیار ما قرار می‌دهند، دانش چندانی نداریم. یک شرکت نوپای کانادایی با نام Genomics Deep، به‌تازگی استفاده از الگوریتم‌های هوش مصنوعی را برای رمزگشایی از معنای ژنوم آغاز کرده است تا بتواند بهترین روش‌های درمانی را برای یک فرد بر اساس DNA سلولی مختص او، تشخیص دهد. الگوریتم‌های یادگیری ساخت این شرکت، جهش‌ها را بررسی می‌کنند و از نتیجه صدها هزار نمونه جهش‌دیده‌ی دیگر، برای پیش‌بینی یک جهش احتمالی استفاده می‌کند.

درحالی‌که آمار جدید ابتلا به سرطان به میلیون‌ها نفر در سال می‌رسد؛ شیمی‌درمانی و داروها، همواره نتوانسته‌اند در درمان آن موفقیت‌آمیز عمل کنند. از آن‌جا که ریشه بسیاری از بیماری‌های افراد در ارتباط با مسائل ژنتیکی است، درک بهتر آرایش ژنتیکی انسان، برای سال‌هایتمادی مورد توجه متخصصان قرار گرفته‌بود؛ اما متأسفانه به دلیل پیچیدگی و حجم بالای داده‌های مورد نیاز، روند پیشرفت‌ها در این عرصه متوقف گردید. با پیشرفت‌های رخ داده در کاربردهای هوش‌مصنوعی و یادگیری ماشین، پژوهشگران از طریق تعیین توالی

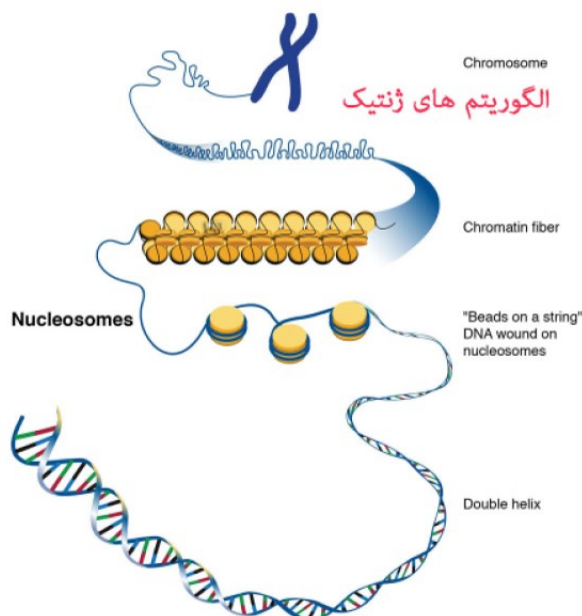
هوش مصنوعی حاصل توسعه سیستم‌های رایانه‌ای است که قادر به انجام وظایفی هستند که به‌طور معمول به هوش انسان نیاز دارند. هوش مصنوعی در بسیاری از زمینه‌ها مانند ژنومیک کاربرد دارد. هوش مصنوعی در زمینه ژنومیک می‌تواند شناسایی هدف داروها و توسعه داروهای جدید را تسهیل کند.

در این مطلب به بررسی چند مورد از کاربردهای هوش مصنوعی در ژنتیک می‌پردازیم:

الف) کارشناسان از نقش پررنگ هوش مصنوعی، در افزایش سرعت و دقت تعیین توالی ژنتیکی و کاهش ریسک خطاهای انسانی سخن می‌گویند.

یکی از حوزه‌هایی که در آن، یادگیری ماشین در حال پیشرفت عظیمی است، مطالعه مجموعه کامل ژن‌ها درون یک ارگانیسم است. در حالی که موضوعاتی نظیر سلامت انسان، توجه زیادی را در این سال‌ها به خود جلب کرده است؛ تعیین توالی ژنتیکی و تجزیه و تحلیل آن نیز می‌تواند انقلابی چشمگیر در عرصه کشاورزی و دامداری ایجاد کند.

پژوهشگران با کمک ابزاری نظیر هوش مصنوعی و به روشی سریع‌تر، ارزان‌تر و دقیق‌تر خواهند توانست توالی DNA را تعیین کرده و آن را تحلیل کنند و در نتیجه، می‌توانند دیدگاهی بهتر نسبت به طرح‌های ژنتیکی خاص به‌دست آورند. با این بینش، آن‌ها قادر خواهند بود در مورد مراقبت از موجوداتی که ممکن است در آینده آسیب پذیرتر باشد یا جهش‌های ژنتیکی که ممکن است موجب بروز بیماری‌های مختلفی شوند و راه‌های مقابله با آن تصمیم‌گیری کنند.



ژنتیکی و ویرایش ژن‌ها، می‌توانند داده‌های ژنومی را بهتر تفسیر کرده و نهایتاً در مورد داروها و درمان‌های مختص به فرد آن‌ها تصمیم‌گیری کنند.

پ) داروها و درمان‌های مختص به فرد:

یکی از جالب‌ترین جنبه‌های فناوری ژنتیک، توسعه پزشکی شخصی است. این حوزه، خدمات پزشکی مختص به یک بیمار یا جمعیتی از افراد با ساختار ژنتیکی مشابه را امکان‌پذیر می‌سازد و پیش‌بینی می‌شود که تا سال ۲۰۳۰، درآمد آن به حدود ۸۷ میلیارد دلار برسد.

در دوران گذشته، هزینه و تکنولوژی از عوامل محدودکننده در پیادسازی پزشکی شخصی محسوب می‌شد؛ اما تکنیک‌های یادگیری ماشین، به غلبه بر این موانع کمک خواهند کرد. ماشین‌ها به شناسایی الگوها در مجموعه داده‌های ژنتیکی کمک می‌کنند و پس از آن، مدل‌های رایانه‌ای می‌توانند درباره احتمال وقوع یک بیماری یا واکنش به تداخلات دارویی در مورد افراد، پیش‌بینی لازم را انجام دهند.

ت) کشف ذهن با هوش مصنوعی:

تقریباً از همان سال ۱۹۵۰ که مقاله‌ی معروف آلن تورینگ (ماشین‌آلات و هوشمندسازی محاسبات) در مجله **Mind** منتشر شد، متخصصان کامپیوتری با استفاده از کاربرد هوش مصنوعی در ژنتیک، به دنبال رمزگذاری ذهن انسان بودند. ذهن، از نظر تئوری، زیرساختی مستقل است؛ بدین معنی که توانایی پردازش آن لزوماً بسته به خیس‌افزار مغز نیست. ذهن را می‌توان روی کامپیوتر بارگذاری کرد یا حتی ذهنی کاملاً جدید در دنیای نرم‌افزاری بازآفرینی کرد.

همه‌ی این مطالب را قبلاً هم شنیده‌ایم. با این‌که هنوز موفق نشده‌ایم ذهنی در قالب نرم‌افزاری بسازیم یا بازسازی کنیم (به جز شبکه‌های عصبی که بازآفرینی ضعیفی از ذهن هستند)، بسیاری از متخصصان کامپیوتر در حال حاضر مشغول کار روی این مسئله هستند.

یکی از کارهای اخیر این حوزه توسط پژوهشگران دانشگاه تارتو (استونی) و دانشگاه پاریس-سکلی (فرانسه) انجام شده است.

این پژوهشگران به جای بازسازی تقریبی ذهن در نرم‌افزار، سوآلی جدید مطرح کردند: آیا می‌توان برای تولید رمز ژنتیکی افرادی که تا به حال وجود نداشته‌اند، از یک الگوریتم استفاده کرد؟ آیا می‌توان در راستای کار تورینگ، از فناوری **GAN** (شبکه‌های مولد تخصصی) استفاده کرد که مدل‌های هوش مصنوعی همچون **BigSleep** را قادر می‌سازد به جای تولید تصاویر واقع‌نگرانه، **DNA** مصنوعی تولید کنند که از یک انسان واقعی قابل تمیز نباشد.

ث) داده‌ها و حریم شخصی:

یلمن می‌گوید: «داده‌های بسیار زیادی در بانک‌های اطلاعات پزشکی وجود دارند که در حال افزایش نیز هستند. با این حال، داده‌های ژنومی

حساسیت بالایی دارند و دسترسی به این بانک‌ها برای پژوهشگران، به دلیل ملاحظات اخلاقی، می‌تواند کار دشواری باشد. هدف اصلی کار ما تولید جایگزینی باکیفیت برای بانک‌های ژنومی و ارائه‌ی راهکاری برای مشکل دسترسی‌پذیری (در قالب یک چارچوب امن اخلاقی) است. لازم است بدانیم مطالعات ما اولین گام این مسیر به شمار می‌روند و هنوز راه زیادی در پیش است.»

جی اضافه می‌کند: «سؤال زیربنایی مطالعات ما این بود که آیا تولید ژنوم‌های مصنوعی به جای ژنوم‌های واقعی می‌تواند به حفظ حریم خصوصی اهداکنندگان ژنوم کمک و در عین حال اطلاعات مفیدی برای جامعه‌ی علوم ژنتیک جمعیت فراهم کند؟»

ج) کاربردها استفاده از ژنوم‌های مصنوعی:

کاربردهای احتمالی ژنوم‌های مصنوعی بازه‌ی گسترده‌ای را در برمی‌گیرند، از درک بهتر تاریخچه‌ی تحولی انسان‌ها گرفته تا ارائه‌ی اطلاعات و بینش به حوزه‌ی پزشکی ژنتیک (از جمله افزایش گوناگونی در این حوزه می‌باشد).

ج) غربالگری:

هوش مصنوعی و یادگیری ماشین کمک می‌کند تا روش‌های اصلاح ژنتیکی دقیق‌تر، ارزان‌تر و آسان‌تر انجام شوند. انتظار می‌رود که در آینده تکنولوژی هوش مصنوعی دربرگیرنده فارماکوژنومیک، ابزارهای غربالگری ژنتیک برای نوزادان و ارتقای روش‌های زراعت و مواردی از این قبیل باشد. با توجه به اینکه ما هنوز توانایی پیش‌بینی آینده را نداریم ولی یک چیز قطعی است: هوش مصنوعی فهم ما را در رابطه با آرایش ژنتیکی خود و دیگر موجودات زنده توسعه خواهد داد.

برخی شرکت‌ها روی فناوری‌هایی کار می‌کنند که با تغییر در سطح سلولی اقدام به ویرایش ژن‌ها می‌کنند. کریسپر، یک تکنولوژی ویرایش ژن است که حاصل تلاش مشترک دانشمندان علوم رایانه و زیست‌شناسی می‌باشد.

هم‌اکنون نتایج مثبتی در عقیم‌سازی ژن‌های عامل بیماری یا اصلاح ژن‌هایی با توانایی تولید محصولات پر بازده صورت گرفته است، ولی همچنان چالش‌های اخلاقی و قانونی در این مبحث مطرح است. اما بیشتر مردم تنها مزایای این گونه اصلاحات ژنتیکی را می‌بینند و تنها زمانی به پیچیدگی این مسئله پی خواهیم برد که روند این اصلاحات در نژاد بشر شروع شود. همچنین مسئله دیگری که متخصصان در روند اصلاح ژنتیکی روی آن کار می‌کنند این است که چگونه باید از هدف‌گیری‌های اشتباه مانند مواردی که به علت شباهت ظاهری دو ژن متخصصان سهوا روی یک ژن اشتباه کار می‌کنند پیشگیری کرد.

مطمئناً در سال‌های آینده با توسعه هوش مصنوعی زندگی بشر با فناوری ماشینی ادغام خواهد شد، سطح زندگی بهبود یافته و عمر بشر طولانی‌تر خواهد شد.

الگوریتم

A*

g = هزینه حرکت برای حرکت از نقطه شروع به یک مربع معین در شبکه می‌باشد.

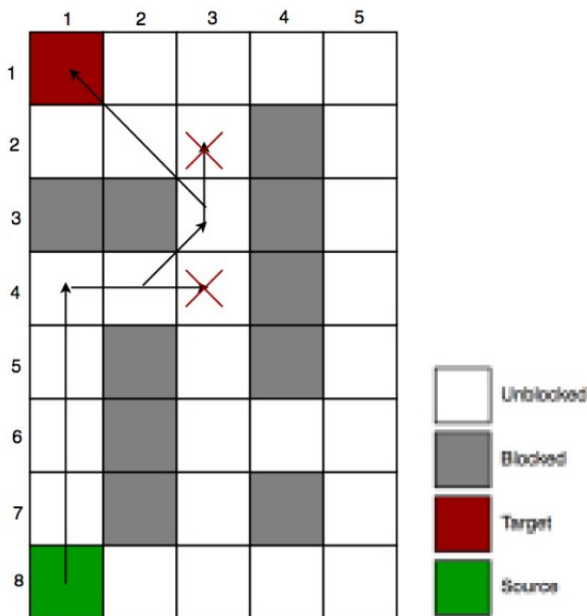
h = هزینه حرکت تخمینی برای حرکت از آن مربع داده شده در شبکه به مقصد نهایی است.

پیدا کردن گره مد نظر اغلب به صورت حریصانه می‌باشد، که چیزی جز نوعی حدس هوشمندانه نیست. چرا که ما واقعاً تا زمانی که مسیر را پیدا نکنیم، فاصله واقعی و موانع پیش‌رو را نمی‌دانیم، پس باید با به کارگیری راه‌ها متعدد مقادیر مد نظر خود را پیدا کنیم.

الگوریتم:

ابتدا دو لیست ایجاد می‌کنیم لیست باز و لیست بسته (درست مانند الگوریتم (Dijkstra)

حال فرض کنید در شکل زیر اگر بخواهیم از سلول مبدا به سلول هدف برسیم، این الگوریتم مسیری را که در زیر نشان داده شده است دنبال می‌کند. توجه داشته باشید که شکل زیر با در نظر گرفتن فاصله اقلیدسی به عنوان یک حالت حریصانه ساخته شده است.



الگوریتم A^* (تلفظ آن A Star) است، الگوریتمی است که به طور گسترده برای مسیریابی (Pathfinding) و پیمایش گراف (Graph Traversal) مورد استفاده قرار می‌گیرد. پیمایش گراف، فرایند پیدا کردن مسیر بین نقاط گوناگونی است که به آن‌ها گره (Node) گفته می‌شود. این الگوریتم می‌تواند برای تقریب کوتاه‌ترین مسیر در موقعیت‌های واقعی، مانند نقشه‌ها به ما کمک کند.

این الگوریتم هوشمند (به اصطلاح دارای مغز) یکی از بهترین و محبوب‌ترین تکنیک‌های مورد استفاده در مسیریابی و پیمایش نمودار است که در تقریب کوتاه‌ترین مسیر به ما کمک می‌کند.

همچنین قابل ذکر است که بسیاری از بازی‌ها و نقشه‌های تحت اینترنت از این الگوریتم برای یافتن کوتاه‌ترین مسیر تقریبی (کارآمد) استفاده می‌کنند.

برای بررسی این الگوریتم باید یک شبکه مربعی با موانع زیادی را در نظر بگیریم که در آن به ما یک سلول شروع و یک سلول هدف داده می‌شود. ما می‌خواهیم در سریع‌ترین زمان ممکن در صورت امکان از سلول شروع به سلول هدف برسیم. در اینجا الگوریتم جستجوی A^* به کمک می‌آید.

کاری که الگوریتم جستجوی A^* انجام می‌دهد این است که در هر مرحله گره را بر اساس مقدار « f » انتخاب می‌کند که پارامتری برابر با مجموع دو پارامتر دیگر « g » و « h » است. در هر مرحله، گره یا سلولی که کم‌ترین « f » را دارد انتخاب می‌کند و آن گره را پردازش می‌کند.

ما « g » و « h » را به ساده‌ترین شکل ممکن تعریف می‌کنیم:

با توجه به داده‌های مسئله ما می‌توانیم g را محاسبه کنیم اما چگونه بهترین h را محاسبه کنیم؟

برای محاسبه آن می‌توانیم از راه‌های زیر استفاده کنیم:

(الف) یا مقدار دقیق h را محاسبه کنید (که مطمئناً زمان بر است).

(ب) مقدار h را با استفاده از برخی راه‌حل‌های حریصانه تخمین بزنیم. (زمان کمتری مصرف می‌شود).

در مورد هر دو روش بحث خواهیم کرد.

الف) کشف مقدار دقیق:

ما می‌توانیم مقادیر دقیق h را پیدا کنیم، اما این معمولاً زمان‌بر است. در زیر برخی از روش‌های محاسبه مقدار دقیق h آورده شده است.

(۱) باید قبل از اجرای الگوریتم، فاصله بین هر جفت سلول را از قبل محاسبه کنید.

(۲) اگر سلول یا موانع مسدودی وجود نداشته باشد، می‌توانیم مقدار دقیق h را بدون هیچ پیش‌محاسبه‌ای با استفاده از فرمول فاصله اقلیدسی پیدا کنیم.

(ب) کشف مقدار تقریبی:

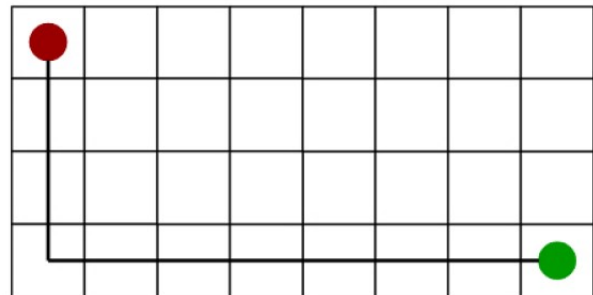
به طور کلی سه اکتشافی تقریبی برای محاسبه h وجود دارد:

(۱) فاصله منتهن:

مجموع قدر مطلق تفاوت‌ها در مختصات سلول هدف و مختصات سلول فعلی، یعنی:

$$h = \text{abs}(\text{current_cell}, x - \text{goal}, x) + \text{abs}(\text{current_cell}, y - \text{goal}, y)$$

چه زمانی از این نوع الگوریتم استفاده کنیم؟ زمانی که فقط در چهار جهت (راست، چپ، بالا، پایین) مجاز به حرکت هستیم.

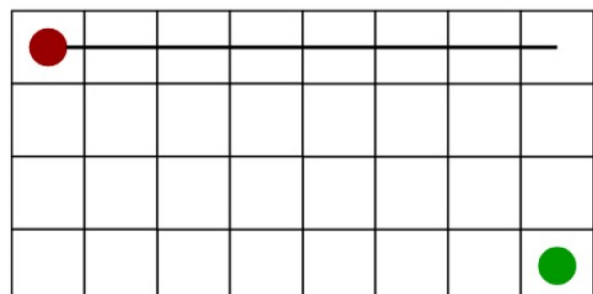


(۲) فاصله مورب: این چیزی نیست جز حداکثر مقادیر تفاوت در مختصات

هدف و مختصات سلول فعلی به ترتیب، یعنی:

$$h = \text{sqrt}((\text{current_cell}, x - \text{goal}, x)^2 + (\text{current_cell}, y - \text{goal}, y)^2)$$

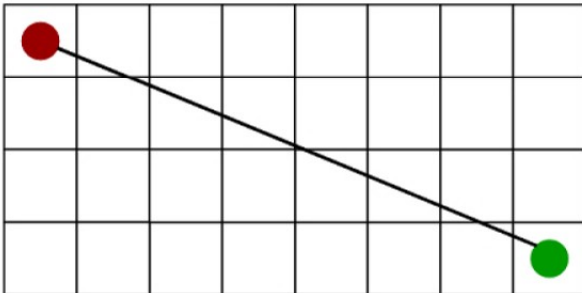
چه زمانی از این اکتشافی استفاده کنیم؟ زمانی که ما اجازه داریم فقط در هشت جهت حرکت کنیم (مشابه حرکت یک پادشاه در شطرنج).



۳) فاصله ی اقلیدسی:

همانطور که از نام آن مشخص است، با استفاده از فرمول، فاصله چیزی جز فاصله بین سلول فعلی و سلول هدف نیست.

چه زمانی از این اکتشافی استفاده کنیم؟ زمانی که به ما اجازه داده شود در هر جهتی حرکت کنیم.



رابطه (شباهت و تفاوت) با سایر الگوریتم‌ها:

Dijkstra یک مورد خاص از این الگوریتم جستجو است که در آن h برای همه گره‌ها برابر صفر است.

پیاده‌سازی:

ما می‌توانیم از هر ساختار داده‌ای برای پیاده‌سازی فهرست باز و فهرست بسته استفاده کنیم، اما برای بهترین عملکرد، از ساختار داده مجموعه‌ای از STL C++ (که به‌عنوان درخت قرمز-سیاه پیاده‌سازی می‌شود) و یک جدول هش بولی برای یک لیست بسته استفاده می‌کنیم. پیاده‌سازی‌ها مشابه الگوریتم Dijkstra است. اگر از پشته فیبوناچی برای پیاده‌سازی لیست باز به جای درخت دوتایی heap یا خود متعادل‌کننده استفاده کنیم، عملکرد بهتر می‌شود (زیرا heap فیبوناچی به طور متوسط زمان $O(1)$ برای درج در لیست باز و کاهش کلید نیاز دارد).

دروس و ابزارها

در هوش مصنوعی

- دروس:

بی‌شک دروس برنامه‌ریزی در هوش مصنوعی، یادگیری ماشین، پردازش تصویر، فرآیندهای تصادفی، پردازش گفتار، پردازش زبان طبیعی و پردازش علائم دیجیتال از مهم‌ترین دروس رشته‌ی هوش مصنوعی هستند.

- زبان برنامه‌نویسی: از پرکاربردترین زبان‌های برنامه‌نویسی برای هوش مصنوعی می‌توان به پایتون اشاره کرد و در کنار آن زبان‌های برنامه‌نویسی لیسپ و پرولوگ (Lisp, Prolog) نیز از مهم‌ترین زبان‌های مورد استفاده در هوش مصنوعی هستند زیرا خصوصیات نحوی و معنایی آن‌ها باعث شده که شیوه‌ها و راه‌حل‌های قوی برای حل مسئله ارائه کنند.

معروف‌ترین کتابخانه‌های هوش مصنوعی پایتون عبارتند از:

برای یادگیری ماشین کار با داده‌ها: TensorFlow

برای تعلیم مدل‌های یادگیری ماشین: Scikit-learn

برای بینایی ماشین و پردازش زبان طبیعی: PyTorch

رابط برنامه‌نویسی برای محاسبات پیچیده‌ی ریاضی: Keras



مسیر یادگیری هوش مصنوعی

و مباحث مقدماتی

ریاضیات و فرمت داده‌ها

مقدمات جبر خطی
آشنایی با توزیع‌های احتمال
آشنایی با پایگاه داده
تفاوت BI و Analytics و Reporting
آشنایی با فرمت‌های مهم داده‌ها
آشنایی با RegEx

برنامه‌نویسی پایتون

دوره مقدماتی آموزش پایتون
آشنایی با کتابخانه NumPy
آشنایی با کتابخانه pandas
آشنایی با Virtual Environment
آشنایی با Jupyter Notebook و
Google Colab
هوش مصنوعی چیست و چه کاربردی دارد؟

منابع جمع‌آوری داده

داده‌کاوی
Web scraping
دیتاست‌های موجود در اینترنت
آشنایی با کگل (kaggle)

کاوش در داده‌ها (EDA)

پیش‌پردازش داده‌ها و کاوش در داده‌ها
کاهش بعد با PCA

جهت دسترسی به نقشه یادگیری ادامه‌ی راه به [این لینک](#) مراجعه کنید.

در ادامه باید تصمیم بگیرید که می‌خواهید در کدام یک از دو رشته زیر

فعالیت کنید: ۱. دانشمند داده ۲. مهندس داده

معرفی کتاب و دوره

در هوش مصنوعی

کتابها:

Deep learning; Authors: Ian Goodfellow, Yoshua Bengio, Aaron Courville

در این کتاب مقدمه ای بر طیف گسترده ای از موضوعات در یادگیری عمیق، که زمینه های ریاضی و مفهومی، تکنیک های یادگیری عمیق مورد استفاده در صنعت و دیدگاه های پژوهشی را پوشش می دهد.

Paradigms of Artificial Intelligence Programming: Case Studies in Common Lisp Author: Peter Norvig

اولین کتابی است که تکنیک های پیشرفته Common Lisp را در زمینه ساختن سیستم های اصلی هوش مصنوعی، با بازسازی برنامه های پیچیده و معتبر هوش مصنوعی با استفاده از پیشرفته ترین Common Lisp آموزش می دهد. این کتاب به دانش آموزان و متخصصان می آموزد که چگونه برنامه های عملی قوی بسازند و اشکال زدایی کنند، در حالی که سبک برنامه نویسی برتر و مفاهیم مهم هوش مصنوعی را نشان می دهد. نویسنده به شدت بر مسائل عملکرد عملی مربوط به نوشتن برنامه های کاری واقعی با اندازه قابل توجه تأکید می کند. فصول مربوط به عیب یابی و کارایی همراه با بحث در مورد اصول برنامه نویسی شی گرا و شرح عملکردهای اصلی CLOS گنجانده شده است.

Basics of Artificial Intelligence & Machine Learning

Author: Dr. Dheeraj Mehrotra

مفهوم هوش مصنوعی و یادگیری ماشین برای سالها با ظهور پیشرفت

تکنولوژی در عمل بوده است. با گذشت زمان، تقریباً هر روایتی از یادگیری، آموزش و لذت بردن، زندگی ما را در هم آمیخته است. در این کتاب هم با توجه به اهمیت این موضوع به بررسی نکات پایهای در این علوم مهم می پردازیم.

برخی از دوره های برخی که می توان در این زمینه معرفی کرد:

[Machine Learning offered by Stanford](#)

این دوره مقدمه ای گسترده برای یادگیری ماشین، داده کاوی و تشخیص الگوهای آماری ارائه می دهد. همچنین از مطالعات موردی و برنامه های کاربردی متعددی استفاده می کند، به طوری که شما یاد خواهید گرفت چگونه الگوریتم های یادگیری را برای ساخت ربات های هوشمند، درک متن، بینایی کامپیوتر و انفورماتیک پزشکی به کار ببرید. این موضوع غیر قابل انکار است که امروزه هوش مصنوعی به طور تصاعدی در حال رشد است.

[Deep Learning A-Z™: Hands-On Artificial Neural Networks](#)

اما هرچه هوش مصنوعی پیشتر پیشرفت کند، مشکلاتی که باید حل کند پیچیده تر می شود و تنها یادگیری عمیق می تواند چنین مشکلات پیچیده ای را حل کند. به همین دلیل است که یادگیری عمیق در قلب هوش مصنوعی قرار دارد و نیاز به یادگیری آن روز به روز در حال افزایش است.

دوره تخصصی یادگیری عمیق به شما کمک می کند قابلیت ها، چالش ها و پیامدهای یادگیری عمیق را درک کنید و همچنین شما را برای مشارکت در توسعه فناوری پیشرفته هوش مصنوعی آماده می کند.

پیانفورماتیک



0100

بیوانفورماتیک یا زیست داده‌موری!

و علوم زیستی را وارد یک بعد جدید از طبقه بندی داده کرده است. اصطلاح بیوانفورماتیک در ابتدا توسط «پائولین هاگگ» و «بن هسپر» در سال ۱۹۷۰ برای توصیف «مطالعه در زمینه پردازش اطلاعات در سیستم‌های زنده» ابداع شد و زمانی که اولین داده‌های توالی بیولوژیکی به اشتراک گذاشته شدند، این اصطلاح هم رایج شد و اولین کاربردهایش را پیدا کرد.

پس از اینکه توالی انسولین در اوایل ۱۹۵۰ توسط «فریدریک سانگر» مشخص شد، دانشمندان به این نتیجه رسیدند که برای ذخیره، سازماندهی، فهرست‌بندی، تفسیر داده‌ها و اطلاعات پیچیده زیستی نیاز به ابزاری دقیق‌تر و سریع‌تر از یک انسان دارند. بدین ترتیب پای کامپیوتر به آزمایشگاه‌ها باز شد و یک دانش میان رشته‌ای به نام بیوانفورماتیک شکل گرفت. این رشته یک مرحله جدید به تست‌های زیست‌شناسی، پزشکی و داروسازی اضافه



بیوانفورماتیک تشکیل شده از دو واژه «بیو» و «انفورماتیک» است که بیو به معنی همه علوم زیستی و انفورماتیک در این جا به معنی علوم ریاضی و فنی است که بر اساس آن هر ایده زیستی را می توان با انجام کار محاسبه‌ای ریاضی و کامپیوتری مورد بررسی قرار داد. به طور کلی علوم زیست یک موضوع بی‌نظم با اطلاعات بسیار زیاد است که بیوانفورماتیک را می‌توان هنر نظم دادن به این اطلاعات دانست.

بیوانفورماتیک یکی از گرایش‌های هوش مصنوعی در مقطع ارشد است اما یک بخش اساسی این رشته را زیست‌شناسی تشکیل می‌دهد. این رشته فقط در مقطع ارشد و دکتری ارائه می‌شود و از گرایش‌های مهندسی نرم افزار و علوم کامپیوتر است. منابع این رشته با منابع گرایش نرم افزار یکی هست. دانشجویان این رشته، در ترم اول دروس زیستی که پیش‌نیاز هستند را فرا می‌گیرند و در ترم‌های بعدی دروس مربوط به یادگیری ماشین و آمار تدریس می‌شود که اساتید عموماً برای تمرین‌ها از داده‌های زیستی استفاده می‌کنند. در طول این مقطع دانشجویان با داده‌های مختلف و طرز برخورد با آن‌ها آشنا می‌شوند. در واقع در این رشته دانشجویان، دانشمندان داده‌ای هستند که به طور تخصصی روی داده‌های زیستی کار می‌کنند.

در ایران اولین دانشجویی که از گرایش «سیستم بیولوژی» رشته «بیوانفورماتیک» از دانشگاه تهران فارغ‌التحصیل شد. «غلامرضا بیدخوری» که در سال ۸۶ در مقطع دکتری بیوانفورماتیک در گرایش سیستم بیولوژی دانشگاه تهران پذیرفته شده بود، با دفاع از پایان‌نامه خود با عنوان «مدل‌سازی پیشروی تومور با استفاده از سیستم بیولوژی» به عنوان اولین دانشجوی دکتری این رشته گرایش در کشور فارغ‌التحصیل شد. استاد راهنمای بیدخوری در طول تحصیل دوره دکتری، دکتر «علی مسعودی‌نژاد»، عضو هیات علمی دانشگاه تهران و پایه‌گذار آزمایشگاه سیستم بیولوژی و بیوانفورماتیک است. رشته بیوانفورماتیک در سال ۱۳۷۵ شمسی برای اولین بار در دانشگاه «بوستون» راه‌اندازی شد که برای اولین بار در سال ۸۳ فرآیند ورود آن به دانشگاه‌های ایران و دانشگاه تهران آغاز شد و «رضا بیدخوری» جزو اولین دانشجویان مقطع دکتری این رشته در دانشگاه تهران است. تعداد شرکت‌های بیوانفورماتیکی در حال حاضر داخل ایران، زیاد نیست اما قطعاً در آینده می‌تواند یکی از شغل‌های پرطرفدار باشد.

در این زمینه «مارگارت دایهوف» پیشگام شد و نخستین پایگاه داده توالی پروتئینی را گردآوری و به‌عنوان کتاب و روش‌های پیشگام در هم‌ترازی توالی و ارزیابی مولکولی منتشر کرد. از جمله افراد دیگر در حوزه شکل‌گیری بیوانفورماتیک «الوین کابات» بود که در تجزیه و تحلیل توالی زیستی در سال ۱۹۷۰ پیش‌قدم شد.

اولین هدف بیوانفورماتیک افزایش درک از فرآیندهای زیستی است؛ به تعبیر دیگر هدف اصلی آن جمع‌آوری و ذخیره‌سازی اطلاعات زیستی، آنالیز و تحلیل آن‌ها برای اکتساب دانش است و تمرکز آن در توسعه و کاربرد تکنیک‌های محاسباتی جامع به منظور کسب این هدف است. در این مورد می‌توان به تشخیص الگو، داده‌کاوی، الگوریتم‌های یادگیری و تصویرسازی اشاره کرد. برای مثال به منظور درک بهتر از چگونگی تغییر فعالیت‌های سلولی نرمال در بیماری‌های مختلف، باید اطلاعات زیستی ترکیب شوند تا تصویری جامع از این فعالیت‌ها شکل بگیرد. از این رو بیوانفورماتیک در زمینه تجزیه و تحلیل و تفسیر انواع مختلفی از داده‌ها به کمک ما می‌آید.

با پیشرفت تکنولوژی و به دنبال آن کسب توانایی خواندن داده‌های زیستی مثل اسید دی‌اکسی‌ریبونوکلئیک؛ مجموعه داده‌ها و اطلاعات بسیار عظیم و پیچیده‌ای در علم زیست‌شناسی کشف شدند که درک و تحلیل آن بدون استفاده از ابزار مناسب تقریباً غیرممکن و بسیار سخت بود. برای مثال مبحث ژنوم (محتوای ژنتیکی یاخته هر فرد؛ در هر یاخته‌ی بدن انسان تمام دستورالعمل‌های کدگذاری شده لازم برای هدایت تمامی فعالیت‌های یاخته و ساخت پروتئین‌های لازم، موجود است. به هر گروه کامل از این دستورالعمل‌ها ژنوم گفته می‌شود)، توالی آن و استخراج اطلاعات ژنومی نیازمند استفاده از ابزارهای محاسباتی پیچیده‌ای است. بیوانفورماتیک انجام این کار را بسیار ساده‌تر از قبل کرده‌است و صرف زمان و هزینه‌ها را به میزان زیادی کاهش داده. این دانش از علوم زیست‌شناسی (به‌خصوص زیست‌شناسی سلولی و مولکولی)، ریاضیات (به‌خصوص آمار و احتمالات) و کامپیوتر در حل و بررسی این مسائل کمک می‌گیرد. به بیان دیگر بیوانفورماتیک با به‌کارگیری تکنیک‌هایی در به‌دست آوردن، ذخیره کردن، تجزیه، تفسیر و تحلیل داده‌های زیستی پیچیده به ما برای حل مشکلات تحقیقات بیولوژیکی کمک می‌کند. این رشته شامل روش‌ها و نرم‌افزارهایی برای درک داده‌های بیولوژیکی و اطلاعات زیستی می‌باشد. رشته

کاربردها و زمینه‌های

بیوانفورماتیک

شما با دانش و تحصیلاتی که در کامپیوتر دارید می‌توانید در نجات جان آدمها و حیوانات و موجودات زنده‌ی روی کره‌ی زمین تاثیر گذار باشید! بیوانفورماتیک با علم هوش مصنوعی در ارتباط است و به کمک هوش مصنوعی بسیار کاربردی‌تر است.

شاید مهم‌ترین و کلی‌ترین کاربردی که بتوان برای بیوانفورماتیک نام برد این است که باعث می‌شود تشخیص علت بیماری‌ها دقیق و سریع‌تر باشد و با انتخاب بهترین پروتکل درمان، جان انسان‌ها را نجات دهد برای مثال در درمان سرطان و بیماری‌های ژنتیکی بسیار موثر است و حتی در این چند سال در تولید واکسن کرونا تاثیر گذار بوده است. البته بیوانفورماتیک فقط در خصوص انسان‌ها نیست بلکه می‌تواند محیط زیست، حیوانات و گیاهان را نیز نجات دهد.

از کاربردهای دیگر بیوانفورماتیک می‌توان به تولید، طراحی، توسعه دارو و درون یارانه‌ای نام برد. درون رایانه‌ای یک شبیه‌ساز کامپیوتری برای سامانه‌ی زیستی است که برای مثال بدن انسان‌ها را با رفتار دقیق ژن‌ها و سلول‌ها شبیه سازی می‌کند؛ آزمایشات و همچنین تست داروها را به جای محیط بدن انسان‌ها در کامپیوتر و یا در شبیه سازی‌های کامپیوتری انجام می‌دهند و به این صورت هزینه‌ها و تلفات را کم می‌کنند.

همچنین از جالبترین کاربردهای بیوانفورماتیک این است که می‌توان به کمک بیوانفورماتیک با توجه به اسید دی اکسی ریبونوکلیک قاتل‌ها و مژنونین، آن‌ها را در صحنه‌های قتل و... شناسایی کرد. برای نمونه‌ای از تشخیص و درمان بیماری‌های ژنتیکی می‌توان به «ماموریت ماسیمو» اشاره کرد؛ ماسیمو کودک ۴ ساله‌ای بود که بیماری «لوکودیستروفی» داشت و برای اولین بار

متخصصان بیوانفورماتیک توانستند این بیماری را تشخیص دهند.

رشته بیوانفورماتیک طیف گسترده‌ای از زیرشاخه‌های بیوتکنولوژیکی را در برمی‌گیرد. بیوانفورماتیک از نظر دامنه کاربرد همچنان در حال رشد است. تعدادی از زمینه‌های که بیوانفورماتیک در آن به کار برده می‌شود عبارت‌اند از:

زیست شناسی محاسباتی - کاربرد راه حل‌های مبتنی بر داده برای مسائل در بیوانفورماتیک.

ژنتیک - مطالعه وراثت و تنوع ویژگی‌های ارثی.

پروتئومیکس - مطالعه پروتئومها و عملکرد آنها.

متازنومیکس - مطالعه ماده ژنتیکی از منابع و نمونه‌های محیطی.

اسید ریبونوئیک ترانسکریپتومیکس - مطالعه رونوشت کامل.

فیلولژنتیک - مطالعه روابط در گروه‌های حیوانات و انسان.

متابولومیک - مطالعه بیوشیمی متابولیسم و متابولیت‌ها.

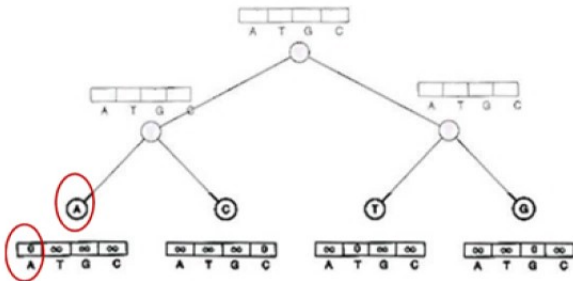
زیست شناسی سیستم‌ها - مدل‌سازی ریاضی و تجزیه و تحلیل مجموعه‌های بزرگ داده‌های زیستی.

تجزیه و تحلیل سازه - ارزیابی که اثرات بارهای فیزیکی بر سازه‌های فیزیکی را تعیین می‌کند.

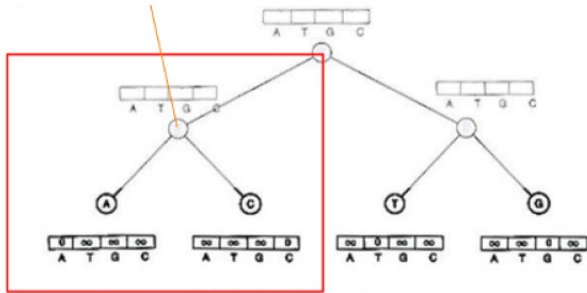
مدل‌سازی مولکولی - مدل‌سازی ساختارهای مولکولی از طریق شیمی محاسباتی.

تجزیه و تحلیل مسیر - یک ارزیابی نرم افزاری که پروتئین‌های مرتبط را در مسیرهای متابولیک شناسایی می‌کند

حل یک مثال از الگوریتم‌های ترکیبیاتی

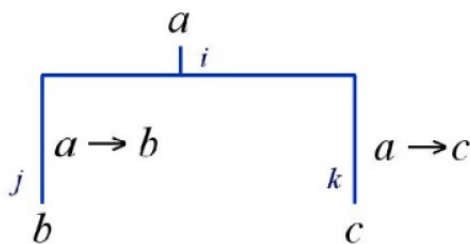


- ابتدا مقدار آرایه این گره را محاسبه می‌کنیم که ریشه زیر درخت سمت چپ است.



- for an internal node i with children j and k :

$$R_i(a) = \min_b (R_j(b) + S(a,b)) + \min_c (R_k(c) + S(a,c))$$

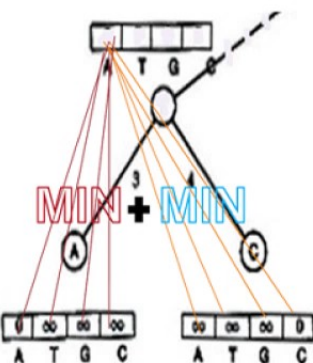


- حال اگر ماتریس امتیازدهی را در نظر بگیریم، داریم:

$R_i(A) = 0 + 9 = 9$

| | $R_j(b)$ | $S(A,b)$ | جمع |
|---|----------|----------|----------|
| A | 0 | 0 | 0 |
| T | ∞ | 3 | ∞ |
| G | ∞ | 4 | ∞ |
| C | ∞ | 9 | ∞ |

| | $R_k(c)$ | $S(A,c)$ | جمع |
|---|----------|----------|----------|
| A | ∞ | 0 | ∞ |
| T | ∞ | 3 | ∞ |
| G | ∞ | 4 | ∞ |
| C | 0 | 9 | 9 |

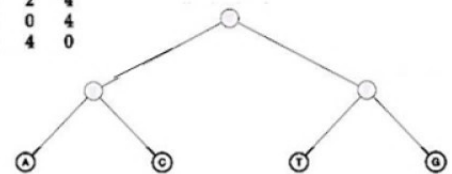


برای آشنایی بیشتر با این شاخه به بررسی و حل یکی از مسائل این حوزه می‌پردازیم، حل این تمرین مربوط به درس «الگوریتم‌های ترکیبیاتی در بیوانفورماتیک» می‌باشد.

سوال:

الگوریتم sank off یا همان پارسیمونی وزن‌دار را بر روی درخت زیر اجرا کنید. (با محاسبه هزینه درخت داده شده از روی ماتریس امتیازدهی، درخت را برچسب‌گذاری کنید).

| δ | A | T | G | C |
|----------|---|---|---|---|
| A | 0 | 3 | 4 | 9 |
| T | 3 | 0 | 2 | 4 |
| G | 4 | 2 | 0 | 4 |
| C | 9 | 4 | 4 | 0 |



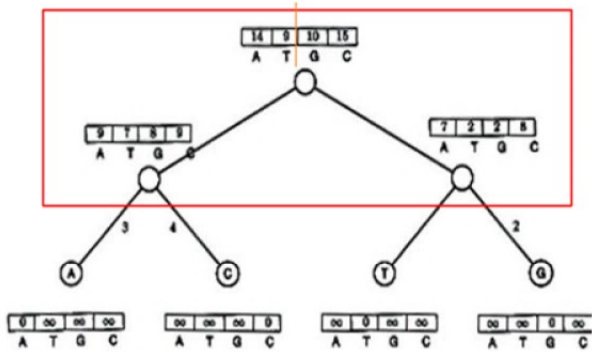
پاسخ:

- برای هر گره آرایه‌ای با 4 درایه تشکیل دهید، به برچسب درایه‌ها دقت کنید.
- (در اینجا ATGC)
- برای محاسبه آرایه مربوط به برگ‌ها عنصری (character) که در برگ قرار دارد درایه‌اش صفر می‌شود و بقیه درایه‌ها بی‌نهایت می‌شوند.

for leaves:

$$R_i(a) = \begin{cases} 0, & \text{if } a \text{ is character at leaf} \\ \infty, & \text{otherwise} \end{cases}$$

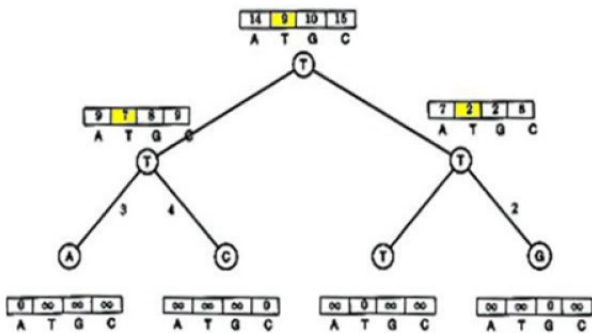
- و در آخر مراحل را برای ریشه با استفاده از مقادیر دو فرزندش محاسبه می‌کنیم.



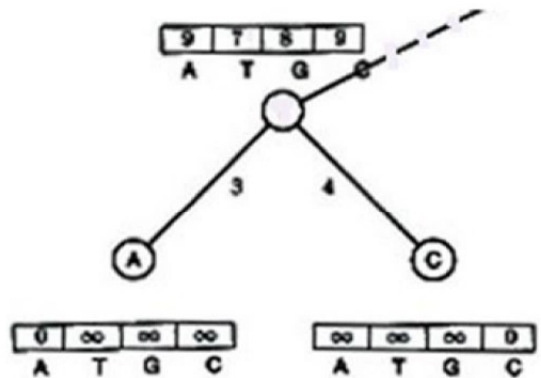
- برای بدست آمدن این مقدار از دو کاراکتر T و T در مرحله قبل استفاده شده پس برچسب T تدها می‌شود.

| δ | A | T | G | C |
|----------|---|---|---|---|
| A | 0 | 3 | 4 | 9 |
| T | 3 | 0 | 2 | 4 |
| G | 4 | 2 | 0 | 4 |
| C | 9 | 4 | 4 | 0 |

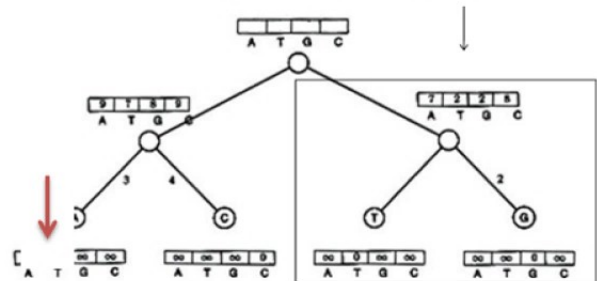
$$9 = (7+0) + (2+0)$$



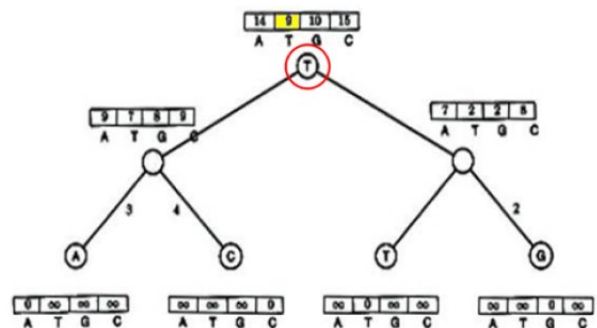
- اگر مرحله قبل را برای T,G,C تکرار کنید، درایه‌های آرایه R_r محاسبه می‌شود.



- مراحل زیر را برای درخت راست ادامه دهید.



- کم‌ترین مقدار آرایه هزینه درخت پارسیمونی است که در اینجا متناظر با کاراکتر T است یعنی T به عنوان برچسب ریشه انتخاب می‌شود.



بررسی بیوانفورماتیک

توالی ژن اینتگراز

ویروس های HIV-1

در بیماران ایرانی

نویسندگان: زهرا حسن‌شاهی، بهرنا دهقانی، طیبه هاشمیپور

بیکتگراویر، دولوتگراویر، الویتگراویر و رالتگراویر هستند. در مطالعات اخیر جهش‌هایی که باعث مقاومت دارویی این پروتئین می‌شوند مورد مطالعه قرار گرفته و چندین جایگاهی به‌عنوان عامل ایجادکننده مقاومت در برابر این داروها تعیین شده‌اند.

در مطالعه حاضر، ۳۵ توالی از ژن اینتگراز بیماران ایرانی مورد بررسی قرار گرفت و نتایج نشان داد که جهش‌هایی که منجر به مقاومت دارویی می‌شوند، در این توالی‌ها وجود ندارند، در حالی که مطالعات نشان داده‌اند که جهش‌های E157Q و R263K باعث شکست درمان می‌شوند و جهش R263K به‌عنوان اولین جهشی که باعث ایجاد مقاومت می‌شود شناخته شده است. همچنین مطالعات اخیر جهش‌های R 118 G و K 263 R را به عنوان جایگاه‌های مقاوم به دارو پیشنهاد دادند که می‌توانند مقاومت ضعیفی در مقابل داروی دولوتگراویر ایجاد کنند. «شیمورا» و همکاران، ۸ جایگاهی آمینواسیدی را برای آنزیم اینتگراز پیشنهاد دادند که در بین آنها ۲ جایگاهی T66I و E92Q باعث مقاومت ویروس به داروی الویتگراویر شد.

«سالادینی» و همکاران به بررسی دقیق جابه‌جایی E157Q پرداختند. بررسی‌های این گروه نشان داد که این جایگاه به‌تنهایی نمی‌تواند عامل ایجادکننده مقاومت دارویی در ویروس باشد. بنابراین به نظر می‌رسد که توالی اینتگراز در بیماران ایرانی دارای جهش مقاومت به درمان نیست و می‌توان با استفاده از داروهای مهارکننده اینتگراز بیماری را کنترل و مهار کرد که البته برای رسیدن به نتیجه قطعی نیاز است که توالی‌های بیشتری مورد بررسی قرار گیرند.

مهارکننده‌های متعددی برای مهار ویروس HIV-1 معرفی شده‌اند ولی به‌علت وجود نمونه‌های مقاوم به درمان اکثر این تلاش‌ها بی‌نتیجه مانده. هدف از مطالعه حاضر نیز بررسی موتاسیون‌های مقاوم به درمان در ژن اینتگراز ویروس ایدز و تاثیر این موتاسیون‌ها بر ساختار، عملکرد و ویژگی‌های فیزیکی و شیمیایی این آنزیم با استفاده از نرم‌افزارهای بیوانفورماتیک می‌باشد.

ویروس HIV-1 از خانواده ویروس‌های رتروویریده است که در بدن انسان لنفوسیت‌های T سیستم ایمنی را مورد حمله قرار می‌دهد و موجب تضعیف آن می‌شود که احتمال ابتلا به عفونت‌های فرصت‌طلب و سرطان‌های متفاوت را زیاد می‌کند در نتیجه باعث مرگ و میر مبتلایان به این بیماری می‌شود.

ویروس HIV-1 از سه ژن اصلی gag، env، pol تشکیل شده است. این ژن‌ها مسئول ساخت پروتئین‌های ساختاری و عملکردی مورد نیاز ویروس هستند. محصولات ژن pol عبارتند از (ترانسکریپتاز معکوس، اینتگراز و پروتاز) که این سه آنزیم برای همانندسازی ویروس ضروری هستند.

آنزیم اینتگراز پروتئینی که وظیفه انتقال ژن ویروس به درون ژن میزبان را دارد. با مهار و کنترل این آنزیم می‌توان مانع ورود ژن ویروس به درون ژن میزبان شد و همچنین روند بیماری را کند و ویروس HIV-1 را مهار کرد، و این نتیجه بررسی‌هایی در ارتباط با این آنزیم انجام شده.

مطالعات اخیر، چندین دارو با عملکردهای مختلف را برای مهار ویروس HIV-1 پیشنهاد داده‌اند که میان آن‌ها داروهای مهارکننده اینتگراز شامل

علاوه بر این، جایگاه‌های اتصال دارو با آنزیم اینتگراز با برنامه Studio Discovery مشخص شد و نتایج نشان داد که داروهای مختلف به جایگاه‌های متفاوتی از آنزیم اینتگراز متصل می‌شوند و بیشترین اتصالات مربوط به ناحیه کاتالیتیک مرکزی آنزیم اینتگراز است. همچنین بیشترین تعداد اتصال به ترتیب بین آنزیم اینتگراز و داروهای دولونتراویر و رالتگراویر بود که منجر به افزایش انرژی اتصالی میان آن‌ها شده است. بنابراین به نظر می‌رسد که دو داروی مذکور بهتر می‌توانند آنزیم اینتگراز را مهار کنند و اثرات درمانی بهتری را داشته باشند.

در مطالعه حاضر، زیرموضوع براساس توالی اینتگراز و با استفاده از سایت standford انجام شد که نتایج نشان داد که زیرنوع تمامی نمونه‌ها A است که با توجه به نقشه ژنی و زیرنوع CRF-35 AD می‌توان نتیجه گرفت که زیرنوع اینتگراز بیماران ایرانی از نوع A است.

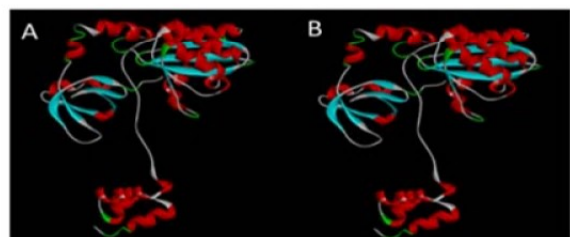
نتایج مطالعه حاضر در زمینه تغییرات پس ترجمه‌ای برای پروتئین‌های اینتگراز و موتانت نشان داد که جایگاه‌های مختلفی تحت تاثیر این فرآیند قرار گرفته‌اند. گلیکوزیلاسیون و فسفوریلاسیون هر دو فرآیندهایی هستند که در بیماری‌زایی و توانایی ویروس در زنده ماندن نقش دارند. با استفاده از نرم‌افزارهای بیوانفورماتیک مشخص شد که شش جایگاه به عنوان جایگاه‌های فسفوریلاسیون در پروتئین اینتگراز قابل پیش‌بینی است که در توالی موتانت، وجود موتاسیون در اطراف جایگاه ۱۷ باعث شد که این جایگاه نیز فسفوریله شود. فسفوریله شدن یک پروتئین مانع از شناسایی آنتی‌ژن‌های آن توسط سیستم ایمنی میزبان می‌شود، در نتیجه به بیماری‌زایی کمک می‌کند و باعث پایداری بیشتر ویروس می‌شود.

در نتیجه جهش‌های موجود در توالی اینتگراز بیماران ایرانی منجر به مقاومت دارویی نسبت به داروهای مهارکننده اینتگراز نمی‌شوند، چراکه جهش‌ها نمی‌توانند بر ساختار سوم این آنزیم و در نتیجه قدرت اتصال دارو با آن تاثیر منفی ایجاد کنند. بنابراین می‌توان پیشنهاد کرد که برای کنترل و مهار بیماران ایرانی آلوده با HIV-1 از داروهای مهارکننده آنزیم اینتگراز به‌خصوص دولونتراویر و رالتگراویر که با قدرت بیشتری به آنزیم متصل می‌شوند، استفاده کرد.

علاوه بر این، برای بررسی دقیق مقاومت دارویی، توالی‌های مورد مطالعه از طریق سایت stanford مورد بررسی قرار گرفتند و طبق نتایج، توالی‌های بیماران ایرانی به داروهای مهارکننده اینتگراز مقاوم نیستند.

آنزیم اینتگراز در ساختار فضایی خود سه دمین انتهایی آمینی، مرکزی کاتالیتیک، انتهایی کربوکسیل دارد. نقش دمین انتهایی آمینی در عملکرد آنزیم به طور دقیق مشخص نیست اما دمین انتهایی کربوکسیل وظیفه اتصال ژنوم ویروس به ژنوم میزبان را دارد و دومین کاتالیتیک مرکزی نیز نقش آنزیمی را ایفا می‌کند. بررسی داده‌ها نشان داد که جهش‌هایی که باعث ایجاد مقاومت دارویی می‌شوند بیشتر در نواحی کاتالیتیک مرکزی قرار دارند، ولی در مطالعه حاضر بیشتر جهش‌های موجود در توالی‌ها در دمین انتهایی آمینی قرار داشت و این دلیلی بر مقاوم نبودن توالی بیماران ایران در برابر داروهای مهارکننده آنزیم اینتگراز بود.

برای بررسی ساختار سوم آنزیم اینتگراز از برنامه ITASSER 5.1 استفاده شد و بعد از بهبود مدل به دست آمده، برهم‌کنش مدل‌های منبع و موتانت با داروها به وسیله برنامه HEX بررسی گردید. نتایج داکینگ نشان داد که جهش‌های موجود در توالی موتانت تفاوت چندانی در قدرت اتصال با داروها ایجاد نمی‌کند و بیشترین اختلاف در قدرت اتصال بین مدل‌های منبع و موتانت با دارو مربوط به برهم‌کنش میان آنزیم اینتگراز و موتانت آن با داروهای دولونتراویر و بیکنگراویر بود که این اختلاف کم می‌تواند به علت وجود جهش در اطراف جایگاه‌های اتصال دارو با آنزیم باشد. علاوه بر این بررسی هم‌ترازی ساختار پروتئین اینتگراز و موتانت نشان داد که این دو پروتئین از نظر فولدینگ و ساختار سوم به میزان ۹۷٪ شباهت دارند (شکل ۱) و وجود جهش‌های موجود در موتانت تاثیر چندانی بر ساختار سه بعدی پروتئین موتانت ندارد و همچنین تاییدی بر نتایج برهم‌کنش آنزیم اینتگراز و موتانت با داروها است، چراکه شباهت ساختار سه بعدی پروتئین، باعث برهم‌کنش‌هایی با انرژی یکسان می‌شود.



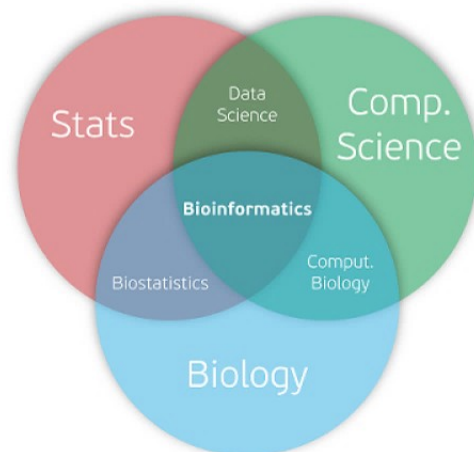
شکل ۱) ساختار سوم پروتئین اینتگراز (A) و موتانت (B) با استفاده از نرم‌افزار I-Tasser

دروس و ابزارها

در بیوانفورماتیک

- یادگیری ماشین در بیوانفورماتیک: با توجه به افزایش حجم و پیچیدگی داده‌ها در بیوانفورماتیک، استفاده از روش‌های نوین تحلیل داده که قابلیت یادگیری و کشف الگوهای موجود در داده را دارند، اجتناب‌ناپذیر است. هدف این درس آشنایی با مفاهیم و روش‌های یادگیری ماشین از نقطه نظر کاربردی و با تاکید روی حل مسائل بیوانفورماتیک است.
- بیوانفورماتیک ساختاری: هدف درس آشنایی با بازیابی، سازماندهی، تجزیه و تحلیل داده‌های زیست‌شناسی ساختاری است. در این درس با روابط توالی، ساختار و عملکرد پروتئین‌ها که از اصول مهم درک جامع بیوانفورماتیک ساختاری است و همچنین روش‌ها، الگوریتم‌ها و نرم‌افزارهای شناخته شده در بیوانفورماتیک ساختاری، آشنا می‌شوید.
- ژنومیک محاسباتی: هدف از این درس، آشنایی با محتوای ژنوم موجودات و نحوه عملکرد آن‌ها می‌باشد. یادگیری مفاهیم ابتدایی تکامل و همچنین پیدا کردن اشتراک‌های مهم بین موجودات از جمله مواردی می‌باشند که در این درس بدان‌ها پرداخته می‌شود. نحوه استخراج و تحلیل دادگان نیز آموزش داده می‌شود. یافتن عوامل تفاوت میان افراد یک گونه و ارتباط میان بیماری‌ها و ژنوم مورد بررسی قرار می‌گیرد.
- الگوریتم بیوانفورماتیک: هدف از این درس آشنایی با مباحث پیشرفته‌تر الگوریتم‌های بیوانفورماتیک است.
- زیست‌شناسی سامانه‌ای محاسباتی: هدف از این درس آشنایی با عملکرد سیستم‌های زیستی بر اساس مدل‌سازی آن با سیستم‌های

بیوانفورماتیک یکی از گرایش‌های هوش مصنوعی در مقطع ارشد است اما یک بخش اساسی این رشته را زیست‌شناسی تشکیل می‌دهد. این رشته فقط در مقطع ارشد و دکتری ارائه می‌شود و از گرایش‌های مهندسی نرم افزار و علوم کامپیوتر است. منابع این رشته با منابع گرایش نرم افزار یکی هست. دانشجویان این رشته، درترم اول دروس زیستی که پیش‌نیاز هستند را فرا می‌گیرند و در ترم‌های بعدی دروس مربوط به یادگیری ماشین و آمار تدریس می‌شود که اساتید عموماً برای تمرین‌ها از داده‌های زیستی استفاده می‌کنند. در طول این مقطع دانشجویان با داده‌های مختلف و طرز برخورد با آن‌ها آشنا می‌شوند. در واقع در این رشته دانشجویان، دانشمندان داده‌ای هستند که به طور تخصصی روی داده‌های زیستی کار می‌کنند. طبق مصوبه مورخ ۰۴/۰۹/۱۳۹۶ دروس اصلی که برای این رشته در دانشگاه شریف ارائه می‌شود به شرح زیر است:



پایتون در حوزه‌های مختلف می‌توان استفاده کرد زیرا پایتون به دلیل دارا بودن ویژگی‌های بسیار خوبی چون کدهای منبع‌باز و کتابخانه‌های وسیع انفورماتیک به محبوبیت بسیاری رسیده است. از دیگر مزایای این زبان برنامه‌نویسی قدرتمند امکانات متفاوت، روان بودن متن، وجود الگوریتم‌های زیستی، آنالیز داده‌های زیستی، در دسترس بودن پایگاه داده‌های زیستی و کاربردی بودن آن است که امکانات بی‌شماری را برای برنامه‌نویسان بیوانفورماتیک فراهم نموده است. همچنین زبان R بسیار کارآمد و به طور گسترده‌ای مورد استفاده قرار می‌گیرد و دارای پایگاه بسیار زیادی از همکاران و کاربرانی است که قابلیت‌های جدید برای تمام جنبه‌های مدرن از تجزیه و تحلیل داده‌ها و تجسم را به آن اضافه می‌کنند؛ علاوه بر آن رایگان و منبع‌باز است. با این حال، با توجه به کاربردها و توانایی‌های R استفاده از آن بدون داشتن پیش‌زمینه و راهنمایی کافی در ابتدا می‌تواند مشکل باشد. به طور کلی می‌توان از زبان‌های برنامه‌نویسی در بیوانفورماتیک از java، Python، R، MySQL، PHP، Perl نام برد.

دینامیکی در سطح مولکولی و سلولی است. به عنوان نمونه‌های اصلی این گونه سیستم‌ها، تمرکز درس بر روی مدل‌های ریاضی بیان ژن و سلول‌های عصبی خواهد بود. این درس دانشجویان را قادر می‌سازد که در تحلیل رفتارهای زیستی در سطح سلولی-مولکولی، از مدل‌های مناسب ریاضی استفاده نموده و به جای توصیف کیفی رفتارهای پایه، آن‌ها را به صورت کمی مدل‌سازی نمایند. آشنایی با ساختارهای شبکه‌ای در تحلیل رفتار سیستم‌های زیستی نیز در این درس، در این راستا صورت می‌پذیرد. زبان‌های برنامه‌نویسی در بیوانفورماتیک: امروزه برنامه‌نویسی در حوزه بیوانفورماتیک یک پدیده نو و رو به پیشرفت است که نظر بسیاری از افراد را جلب نموده‌است. زبان برنامه‌نویسی در بیوانفورماتیک بستگی به حوزه‌ای که کار می‌شود دارد. هر زبانی دارای نقاط قوت و ضعف است. شما باید زبانی را انتخاب کنید که به بهترین وجه برای وظایفی که در دست دارید مناسب است. برای مثال از زبان C می‌توان برای «شبیه‌سازی ملکولی» استفاده کرد و از جاوا برای «تفسیر توالی‌یابی نسل بعدی». همچنین از

معرفی کتاب و دوره

در بیوانفورماتیک

کتاب:

- بیوانفورماتیک کاربردی:

کتاب بیوانفورماتیک کاربردی نوشته «پاول سلزر، ریچارد مارهوفر، الیور کوخ» که در سال ۲۰۱۸ توسط نشریه معتبر و علمی «اشپرینگر» به چاپ رسیده، حاوی اصول و اطلاعات کاربردی ارزشمندی در مورد ابزارهای بیوانفورماتیکی پرکاربرد برای پژوهش و تحقیق در زمینه بیوانفورماتیک می‌باشد و از این حیث منبع مناسبی برای محققین، اساتید و دانشجویان حوزه ژنتیک، پزشکی، دارویی، کشاورزی و علوم زیستی است. این کتاب خوانندگان را با اصول اولیه بیوانفورماتیک و کاربرد عملی و استفاده از ابزارهای محاسباتی، بدون پیش‌فرض قبلی در برنامه نویسی یا انفورماتیک آشنا می‌کند.

- بیوانفورماتیک به زبان ساده:

رشته بیوانفورماتیک برای مبتدیان و کسانی که در این رشته فعالیت نمی‌کنند ممکن است سخت و پیچیده به نظر برسد اما کتاب بیوانفورماتیک به زبان ساده یک نمای تقریباً ساده از دانش بیوانفورماتیک برایتان نمایش می‌دهد. این کتاب نوشته «ژان میشل کلاوری» و «سدریک نوتردام» است که توسط یکی از بزرگترین ناشران علمی جهان یعنی کمپانی بزرگ «Wiley» به چاپ رسیده. احتمالاً با سری کتابهای «for dummies» آشنا هستید هدف این کتابها توضیح علوم پیچیده مانند علوم مربوط به کامپیوتر به زبان ساده برای مبتدیان و افراد نیمه حرفه‌ای است.

معرفی دوره:

آموزش مقدماتی بیوانفورماتیک فرادرس:

آموزش بیوانفورماتیک مقدماتی شامل دو بخش است. در بخش اول مروری بر زیست مولکولی انجام می‌شود و سپس بیوانفورماتیک و تاریخچه آن و در ادامه تطابق دوتایی و چندتایی توالی‌ها، جستجوی توالی، مدل مارکوف مخفی، پروفایل، PSSM و در نهایت ساخت درخت فیلوژنتیکی (Phylogenetic tree) آموزش داده می‌شود. این آموزش هم مفاهیم تئوری و هم عملی می‌پردازد و نرم افزارها و سایت‌ها و دیتابیس‌های اولیه را آموزش می‌دهد.

در بخش دوم زبان R آموزش داده می‌شود. این آموزش مقدماتی است و مفاهیم پایه برای یک دانشجوی علاقه مند به کار در مباحث بیوانفورماتیک آموزش داده می‌شود. همچنین با تعدادی از پکیج‌های بیوانفورماتیکی آشنا می‌شود.

مدرس‌ها:

دکتر فرشید شیرافکن، دکتری تخصصی بیوانفورماتیک

فهرست سرفصل‌ها:

مروری بر زیست مولکولی، مقدمات بیوانفورماتیک، پایگاه داده‌های زیستی، انطباق دوتایی توالی‌ها، جستجوی تشابه در پایگاه داده‌ها، انطباق چندگانه توالی، پیشگویی دمن و موتیف، فیلوژنتیک، آموزش برنامه‌نویسی R

اینترنت اشیا، هویت بخش دیجیتال به اشیای بیجان

استفاده کرد و (IoT) را به عنوان تکنولوژی‌ای که می‌تواند با استفاده از RFID دستگاه‌های زیادی را بهم مرتبط کند معرفی کرد. او در ارائه‌ای که از این موضوع داشت به عمد از واژه‌ی اینترنت استفاده کرد تا توجه دیگران را به این موضوع جلب کند چون در آن زمان اینترنت به شدت موضوع پربحثی بود. با اینکه ایده اشتون از دستگاه‌های مبتنی بر RFID با دستگاه‌های مبتنی بر IP فرق می‌کند اما ایده‌ی او نقش کلیدی در پیشرفت تکنولوژی (IoT) داشت.

همچنین در ابتدای قرن بیستم اینترنت اشیا مورد توجه ویژه‌ی رسانه‌های جهان از جمله گاردین، فوربز و بوستون گلوب قرار گرفته بود. علاقه به اینترنت اشیا به سرعت در حال پیشرفت بود که باعث شد تا اولین کنفرانس جهانی اینترنت اشیا در سال ۲۰۰۸ با شرکت ۲۳ کشور در سوییس برگزار شود.

علاوه بر این‌ها، توسعه‌های اساسی زیادی باعث پرورش (IoT) شدند. مثلاً LG در سال ۲۰۰۰ یخچال متصل به اینترنت را معرفی کرد که اجازه‌ی خرید آنلاین و تماس ویدیویی را به کاربر می‌داد. نمونه‌ای دیگر، رباتی به شکل خرگوش به نام «نابازتاگ» بود که قادر به گفتن آخرین اخبار، وضعیت آب و هوایی و بورس بود؛ که در سال ۲۰۰۵ ساخته شد. حتی در آن زمان‌ها بنابر اطلاعات سیسکو، تعداد دستگاه‌های بیجان به هم مرتبط از تعداد انسان‌ها عبور کرده بود.

بحث دستگاه‌های مرتبط اولین بار در سال ۱۸۳۲ زمانی که اولین دستگاه تلگراف الکترومکانیک طراحی شده بود، مطرح شد. تلگراف امکان ارتباط مستقیم را بین دو دستگاه به وسیله سبک‌ال‌های الکتریکی فراهم می‌کرد. اما داستان اصلی اینترنت اشیا در سال ۱۹۶۰ با اختراع اینترنت شروع شده است.

اولین دستگاه آی‌اوتی (IoT)، دستگاه فروشی (Coca-Cola vending machine) بود که در دانشگاه کارینگ ملون قرار داشت و توسط دانشجویان همان دانشگاه برنامه‌نویسی شده بود. آن‌ها با قرار دادن میکروسویچ‌هایی در این دستگاه و استفاده از ورژن ابتدایی اینترنت، سردی نوشیدنی‌ها و تعداد آن‌ها را بررسی می‌کردند. این اختراع سرعت مطالعات بر روی اینترنت اشیا و توسعه ماشین‌های به هم پیوسته را بالا برد.

در سال ۱۹۹۰ «جان رامکی» با استفاده از پروتکل (TCP/IP) برای اولین بار دستگاه برشته کن (توستر) را به اینترنت متصل کرد.

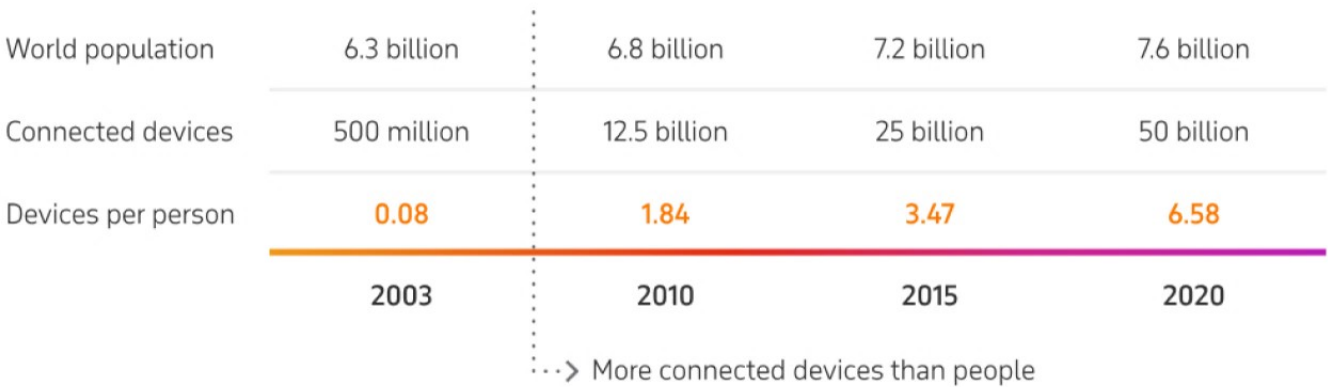
یک سال بعد، دانشمندان دانشگاه کمبریج با استفاده از دوربین و اینترنت مقدار قهوه موجود در ظرف قهوه را تحت نظر داشتند. آن‌ها دوربین را طوری برنامه‌ریزی کرده بودند که در دقیقه سه بار از مقدار قهوه عکس تهیه کند و عکس را به کامپیوتر محلی ارسال کند تا همه از میزان قهوه باخبر باشند.

در سال ۱۹۹۹، «کونین اشتون» برای اولین بار از اصطلاح اینترنت اشیا



Connected devices by 2020

Cisco IBSG, April 2011



Data source: Cisco—The internet of things. How the next evolution of the internet is changing everything

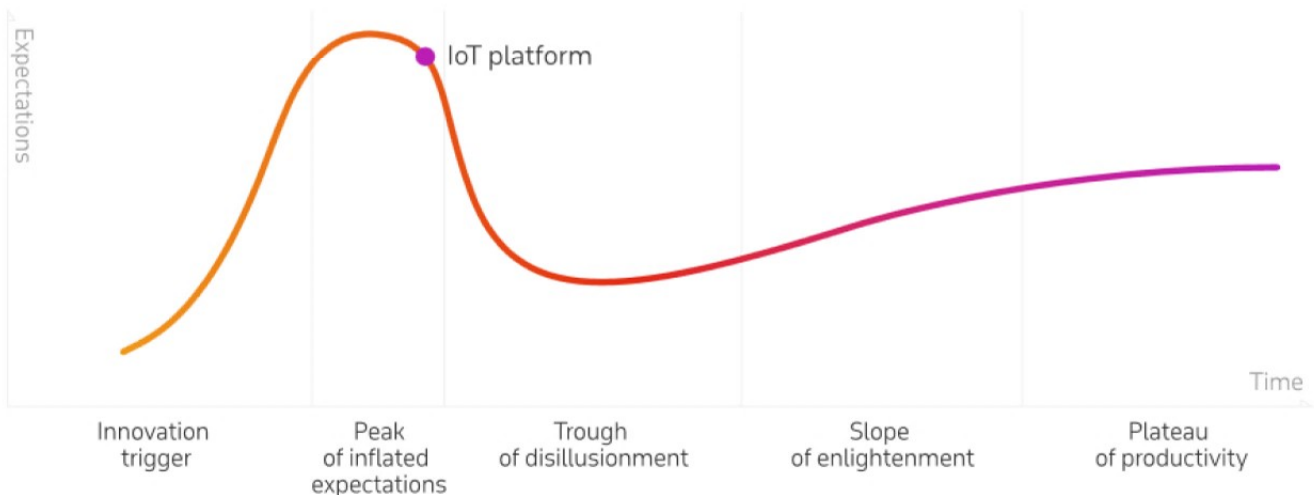
پس در حالت کلی اینترنت اشیا به دنیای دستگاه‌های فیزیکی متصل شده به اینترنت گفته می‌شود. تا قبل از این اینترنت برای مدیریت اشیا مجازی مثل ایمیل، شبکه‌های اجتماعی و تارنماها بود. با ورود اینترنت اشیا، می‌توان با اینترنت، اشیا فیزیکی همانند خودروها، لوازم منزل، ساعت مچی، گوشی‌های هوشمند و... را مدیریت کرد. این اشیا علاوه بر ارتباط با اینترنت، با یکدیگر هم در ارتباط هستند و تبادل اطلاعات می‌کنند.

پیشرفت های (IoT) با پیوستن‌اش به «هایپ گارتتر» (Gartner Hype) در سال ۲۰۱۱ چشمگیر شد. در همان سال IPv6 عمومی شد. از آن زمان تا به حال دستگاه‌های به هم مرتبط به شدت گسترش پیدا کرده و در زندگی روزانه ما جا گرفته‌اند. گوگل، سیسکو و جنرال موتورز تلاش‌شان را بر روی ساخت سنسورها و دستگاه‌ها گذاشته اند. اینترنت اشیا تقریباً در همه‌ی صنایع از جمله در درمان، حمل و نقل، نفت و انرژی، کشاورزی و ... نفوذ پیدا کرده است.

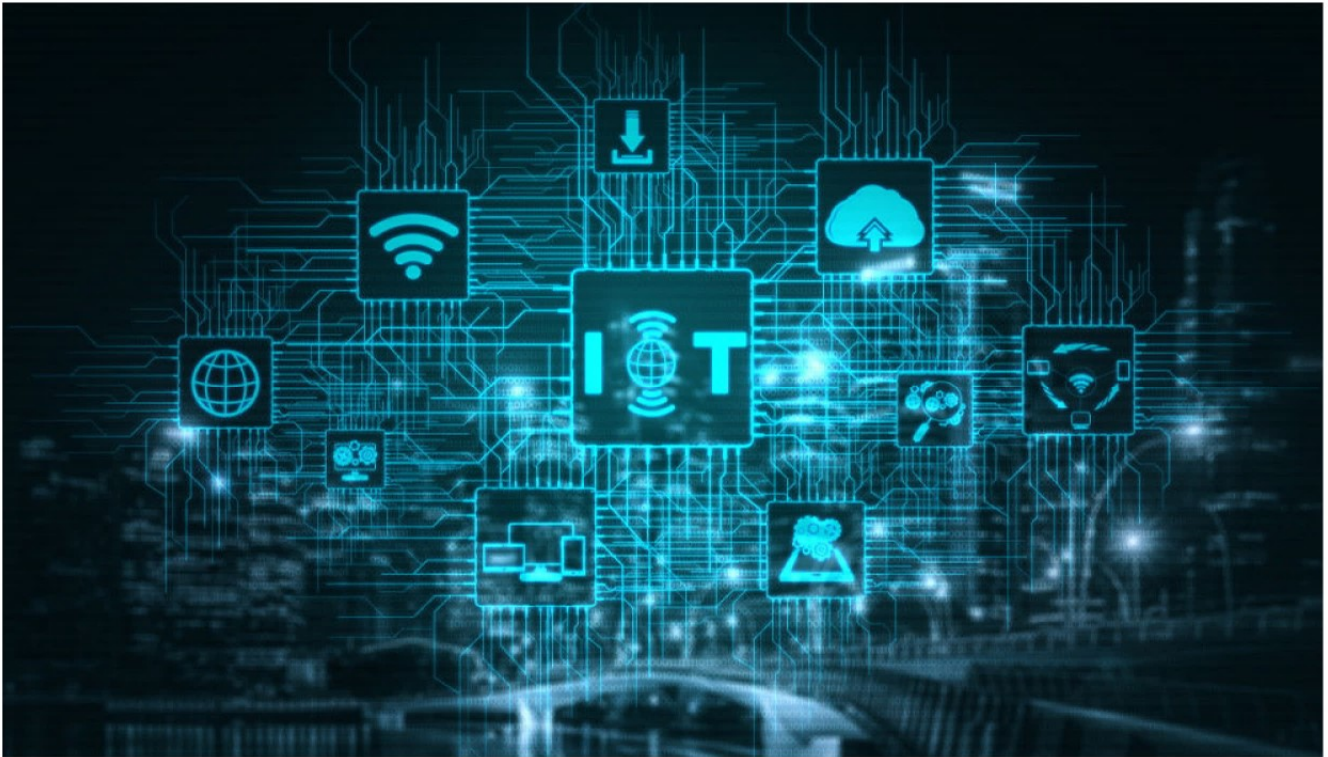
Hype Cycle for emerging technologies, 2018



● Plateau will be reached in 5 to 10 years



Data source: Gartner—5 trends emerge in the Gartner Hype Cycle for emerging technologies



استفاده می‌کنند. از جمله این امکانات عبارتند از: ایجاد مدل‌های جدید تجارت و جریان‌های سود، بهینه‌سازی تصمیمات تجاری از طریق آمارهای به دست آمده از داده‌های تحلیل شده به وسیله IoT، برای افزایش سوددهی و بهره‌وری عملیات تجاری و افزایش تجربه کاربری. همچنین ۴ مورد زیر می‌تواند درحالت کلی اهمیت این رشته را توصیف کند:

- ۱) دریافت اطلاعات بیشتر در زمان واقعی
- ۲) کنترل اشیا از راه دور
- ۳) اتوماتیک کردن همه چیز
- ۴) انجام سریع تر کارها

اینترنت در عصر حاضر همه‌ی مردم را به هم متصل می‌کند اما اینترنت اشیا همه‌ی اشیا را به یکدیگر متصل می‌کند. عبارت اینترنت اشیا جهانی را توصیف می‌کند که در آن هر چیزی از جمله اشیا بی‌جان، برای خود هویت دیجیتال داشته باشند و به کامپیوترها اجازه دهند تا آن‌ها را سازماندهی و مدیریت کنند.

این رشته به دلیل فراهم‌سازی امکان همکاری دنیای فیزیکی با دنیای دیجیتالی حائز اهمیت است و همچنین مزایای فراوانی برای موسسات و شرکت‌های مختلف به‌همراه داشته؛ چراکه برای آن‌ها امکان ساده‌سازی و رسیدگی به اهداف روزانه را فراهم می‌سازد. با رشد روزافزون اینترنت اشیا کمپانی‌های مختلف از امکانات کلیدی‌ای که این رشته به آن‌ها ارائه می‌دهد

جایگاه IoT

در زندگی هوشمند

- مدیریت آلودگی هوای کلان‌شهرها: برای این کار می‌توان بیش از صدها سنسور تشخیص آلودگی هوا را در سطح شهرها قرار داد که همگی آن‌ها به صورت برخط اطلاعات آن منطقه را به مرکز کنترل ارسال کنند. با جمع‌آوری و تحلیل این اطلاعات می‌توان فهمید کدام نقاط شهر در چه مواقعی از روز آلوده‌تر هستند؛ که با برنامه ریزی دقیق می‌توانند هشدارهای لازم را به مردم بدهند و تردد را نیز در آن منطقه کاهش دهند.
- افزایش بهره‌وری کشاورزی: یک سنسور هوشمند را در داخل خاک قرار می‌دهند که این سنسور میزان مواد معدنی و رطوبت و میزان تابش نور را اندازه می‌گیرد. سپس اطلاعات را جمع‌آوری می‌کند و بر اساس آن استانداردهایی که برای دستگاه تعریف شده، مقدار کود و آب تزریقی به خاک و تابش نور را خودش تنظیم می‌کند.

- از کاربردهای اینترنت اشیا می‌توان به موارد زیر اشاره کرد:
- پزشکی: استفاده از فناوری‌های پوشیدنی یا سنسورهای متصل به بیماران، به پزشکان اجازه‌ی تحت نظر گرفتن شرایط بیمار در خارج از بیمارستان و به صورت آنی را می‌دهد. به خاطر تحت نظر داشتن مداوم علائم حیاتی بیمار اینترنت اشیا به بهبود سلامتی بیمار و جلوگیری از اتفاقات منجر به مرگ در بیماران با مخاطره بالا جلوگیری می‌کند. استفاده دیگر از اینترنت اشیا در تخت‌های بیمارستان است که با داشتن سنسورهای فشار خون، علائم حیاتی، درجه بدن بیمار و ... را کنترل می‌کند.
- سیستم تهویه‌ی خانه‌ها: سیستم‌های تهویه هوشمند دمای هوای بیرون از منزل و پیش‌بینی هوای آن روز را به صورت برخط دریافت می‌کنند و بر همین اساس دمای داخل خانه را تنظیم می‌کنند.



چالش‌های حل شده و حل نشده در اینترنت اشیا

علاوه بر کاربردهای گسترده این رشته هنوز هم مشکلات و محدودیت‌هایی در رابطه با برخی مسائل وجود دارد که برخی از این مشکلات حل و برخی هنوز حل نشده باقی مانده‌اند در زیر به برخی از آن‌ها اشاره می‌کنیم:

هر مشکل حل نشده ای در دنیای این علم، در یکی از سه بخش اساسی اینترنت اشیا نهفته است.

۱- نرم افزار:

- یادگیری ماشینی: چگونه می توان یادگیری ماشین را به طور موثر به دستگاه های اینترنت اشیا آورد؟
- امنیت: ایجاد فضای سایبری ایمن برای عملکرد دستگاه های اینترنت اشیا.
- زیرساخت ابری: ساخت زیرساخت ابری بهتر و اختصاصی برای اینترنت اشیا.
- رابط های کاربردی: ایجاد رابط های خاص اینترنت اشیا و نرم افزارهای کاربردی برای داده های بزرگ.
- تجزیه و تحلیل: جمع آوری حجم زیادی از داده ها به طور موثر و سپس استخراج بینش های مفید. ایجاد الگوریتم های بهتر، کارآمد و جدید برای مدیریت حجم زیادی از داده ها نیز یکی از مشکلات حائز اهمیت است.

۲- ارتباط:

- قابلیت تعامل: اینترنت اشیا از استانداردهای ارتباطی مختلفی استفاده می‌کند و برخی از گره‌های حسگر در یک محیط معین ممکن است با استفاده از یک پروتکل خاص در ارتباط باشند (همانند Zigbee)، همچنین ممکن است مجموعه‌ی دیگری از گره‌ها وجود داشته باشند که از آن‌ها استفاده می‌کنند (مانند بلوتوث).

مشکل آن است که بفهمیم گره‌هایی که از پروتکل‌های مختلف استفاده می‌کنند چگونه می‌توانند با هم کار کنند و همچنین تشابه مترجم هم‌زمان برای دو و یا چند زبان را چگونه مدیریت کنیم.

- گره‌های استاتیک و متحرک: یکی دیگر از مشکلات خوب، که ارزش حل دارد، ارتباط کارآمد بین گره‌های حسگر است، زمانی که برخی یا همه‌ی آن‌ها متحرک هستند.
- مقیاس پذیری: چگونه می توان شبکه را گسترش داد تا تعداد زیادی گره را در خود جای دهد؟
- ارتباطات دوربرد: پروتکل‌های ارتباطی که می‌توانند در فواصل دورتر داده‌ها را منتقل و در عین حال انرژی کمتری مصرف کنند.

۳- سخت افزار:

- سنسورها: مشکلات زیادی در اینجا وجود دارد که می توان آنها را حل کرد.
- MEMS (حسگرهای میکرو الکترومکانیکی)، NEMS (سنسورهای نانو الکترومکانیکی)، مدارهای رابط حسگر، حسگرهای مجهز به اینترنت اشیا از جمله مواردی هستند که مشکلات جدید زیادی دارند که عمدتاً در جنبه های طراحی و مواد قابل حل هستند.
- سیستم های جاسازی شده: توسعه و طراحی پردازنده های جاسازی شده ی جدید با قدرت محاسباتی بالاتر، مصرف انرژی کمتر، اندازه کوچکتر و ... از جمله مشکلاتی هستند که به طور مداوم روی آنها کار می شود.
- مدل سازی قدرت دستگاه: پیکربندی سخت افزار برای مصرف انرژی کمتر نیز یکی دیگر از مشکلات موجود است.

حضور یک پرستار در خانه‌ی فرد سالمند نخواهد بود و هزینه‌های بیمار کاهش می‌یابد.

۳. مبارزه با ایزوله‌ی اجتماعی:

سالمندان با بالا رفتن سن تمایل به گوشه‌گیری و در نهایت تنهایی و ناتوانی در حفظ کردن روابط اجتماعی دارند. دستگاه‌های اینترنت اشیا امکان ارتباط سالمندان با افراد دیگر را فراهم می‌کنند، همچنین ربات‌ها هم در این موضوع نقش دارند. علاوه بر موارد بالا دستگاه‌هایی وجود دارند که به نزدیکان فرد درباره‌ی او اطلاعاتی مثل بیرون رفتن و ... را می‌دهند.

۴. تهیه‌ی داده:

دستگاه‌های اینترنت اشیا درباره‌ی شرایط بیمار اطلاعات را دریافت و برای پزشکان ارسال می‌کنند در نتیجه هنگام بروز مشکلات، رسیدگی به بیمار سریع‌تر و با نتایج بهتری خواهد بود. این اطلاعات شامل علائم حیاتی بیمار و یا موقعیت مکانی او است.

یکی از مسائل مهم و کاربردی حل شده در اینترنت اشیا چگونگی کمک به جمعیت سالمندان می‌باشد؛ که به شرح زیر است:

۱. کاهش افتادن‌ها:

دانشمندان با ساخت شتاب‌سنج‌هایی که الگوریتم راه رفتن فرد و نبودن ثبات در شیوه‌ی راه رفتن او را ثبت می‌کنند احتمال افتادن افراد را پیش‌بینی و با خبر دادن به پزشک مربوطه از افتادن او جلوگیری می‌کنند. همچنین مسیر را برای سالمندان شرح داده و از بروز اتفاقات ناگوار دیگر نیز پیشگیری می‌کند.

۲. یادآوری برای مصرف داروها:

اینترنت اشیا به پزشک بیمار اطلاع می‌دهد که او چه زمانی و به چه شکل دارویش را مصرف کرده و آیا طبق دستور بوده است یا خیر. و به همین علت نیاز به مراجعه‌ی مکرر بیمار به بیمارستان نیست. علاوه بر این نیاز به

کاربرد اینترنت اشیا در کنترل بیماری کووید ۱۹

نویسندگان: سمیه سلیمه، ناصر خنجری، سید پیمان عمادی

معماری اینترنت اشیا: حسگرها داده‌ها را دریافت، سپس این داده‌ها در انبار داده‌های بزرگ پالایش شده و با استفاده از ماشین، مدلی برای سیستم‌های تصمیم‌گیرنده براساس نیازها و داده‌ها ایجاد می‌شود. تجزیه و تحلیل داده‌ها برای پیش‌بینی نتایج و مقایسه عملکرد مدل‌ها استفاده می‌شود. در آخر حسگرها از طریق یک یا چند دستگاه دروازه‌ای برای ارتباط برقرار کرده، و نتایج را به ابر منتقل می‌کنند.

در فناوری اینترنت اشیا می‌توان هر شی را شناسایی، آدرس‌پذیر، کنترل و بر آن نظارت کرد. اینترنت اشیا به عنوان انقلابی در فناوری اطلاعات و ارتباطات شناخته شده است و پذیرش آن فرصت‌هایی را برای بسیاری از حوزه‌ها و شرکت‌ها فراهم می‌کند. بیماری همه گیر کووید ۱۹ بسیاری از سازمان‌ها را وادار کرد تا روند کار را از دور دنبال کنند و برای این کار به فناوری‌های پیشرفته روی آورده‌اند.

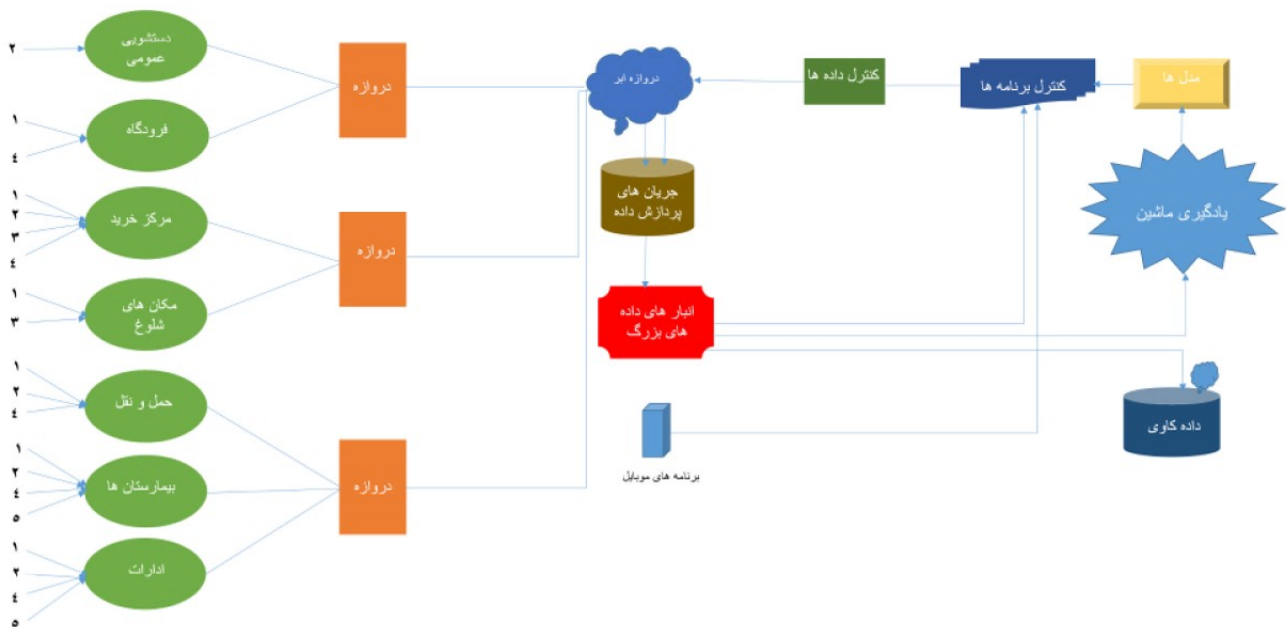
اینترنت اشیا و پزشکی: اینترنت اشیا در فراهم سازی امکان پزشکی از راه دور در دوران همه گیری کووید ۱۹ نقش کلیدی دارد که سیستم‌های مراقبت بهداشتی اینترنت اشیا این امکان را فراهم می‌کنند. اینترنت اشیا امکان در قرنطینه ماندن فرد مبتلا و نظارت بر سلامتی او از راه دور را در دوره همه‌گیری تامین می‌کند.



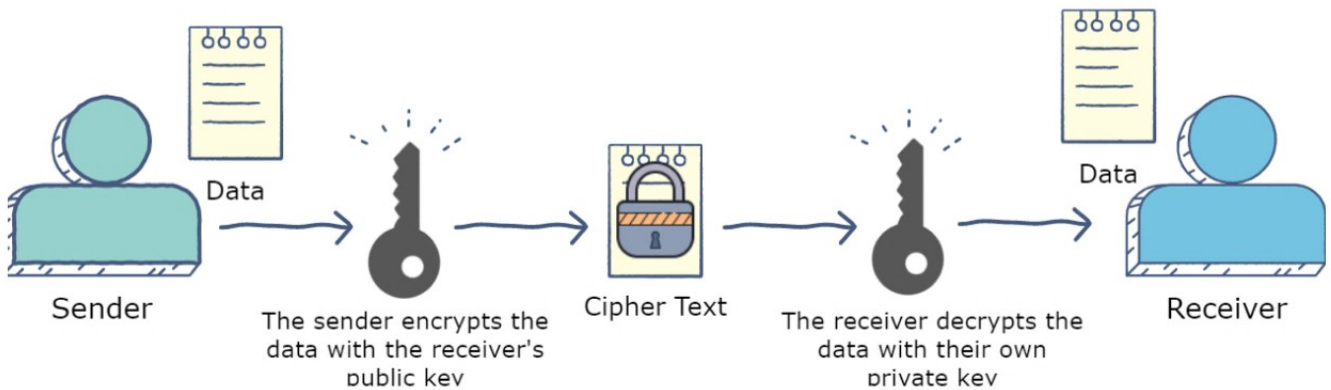
| ردیف | حسگر/دستگاه ها | کاربرد ها |
|------|-------------------|------------------------------------------------------------------|
| ۱ | دماسنج مادون قرمز | اندازه گیری دمای بدن / تصویربرداری حرارتی |
| ۲ | سنسور مادون قرمز | کنترل: عملکرد درب و پنجره ها ، عملکرد آسانسور ، عملکرد سوئیچ برق |
| | | تأمین آب در سرویس بهداشتی |
| ۳ | ساعت هوشمند | تشخیص ضربان قلب |
| ۴ | دوربین نوری | معاینه و تشخیص از روی چهره بیمار |
| ۵ | دوربین تحت شبکه | برای پخش کنفرانس / نشست / خبر مجازی |

- پزشکان بهداشت: پزشکان موارد مشکوک بارگذاری شده توسط کاربران را کنترل و اگر نشانگر عفونت باشد می‌توانند به سرعت به آن پاسخ دهند؛ همچنین اجازه هرگونه تحقیقات بالینی بیشتری برای تایید موارد مشکوک را خواهند داشت.
 - زیرساخت ابر: زیرساخت ابر از طریق اینترنت به هم متصل شده که اجازه‌ی بارگذاری داده توسط کاربر در هر زمان، ذخیره سوابق بهداشتی فرد، پیش بینی نتایج، اشتراک توصیه‌های پزشکی با فرد و ذخیره اطلاعات را فراهم می‌سازد.
- اینترنت اشیا با کاربردهای بسیار خود، تقریباً در تمامی حوزه‌ها باعث بهبود و کارایی عملکرد می‌شود. این فناوری در حوزه بهداشت و درمان بسیار مفید بوده و با ارائه خدمات مراقبت از راه دور، کنترل و نظارت بیماران در هر زمان و مکان، کاهش هزینه‌های جاری و همچنین کاهش هزینه‌های مربوط به بستری و نیز ارائه خدمات سودمند و متنوع دیگر می‌تواند نقش موثر خود را در این حوزه ایفا کند.

- چارچوب پیشنهادی اینترنت اشیا پزشکی، دارای پنج مولفه اصلی است:
- جمع‌آوری اطلاعات و بارگذاری علائم: از طریق حسگرهای پوشیدنی بر روی بدن کاربر علائم بیماری مثل سرفه، تب، خستگی، گلودرد را جمع‌آوری می‌کنیم. همچنین سایر داده‌ها مثل سفر و سابقه تماس را می‌توان از طریق تلفن همراه به‌دست‌آورد.
- مرکز قرنطینه: سوابق داده‌ی کاربرانی که در قرنطینه یا یک مرکز پزشکی تحت نظر بوده‌اند دریافت می‌کند. این سوابق را به دو دسته فنی _ که شامل داده‌های سری زمانی که در مرحله قبل جمع‌آوری شده‌است _ و غیرفنی _ که شامل اطلاعات ۳-۴ هفته گذشته مثل بیماری‌های مزمن، سابقه تماس، و سایر اطلاعات مرتبط با خانواده، سوابق بیماری و... است _ تقسیم بندی می‌کنند.
- مرکز تجزیه و تحلیل داده‌ها: مرکز داده میزبان تجزیه و تحلیل داده‌های مناسب برای یادگیری ماشین است. بر اساس داده‌های به روز جمع‌آوری شده و بارگذاری شده از کاربران می‌توان نتیجه درمانی بیمار را پیش‌بینی کرد. مدل‌های توسعه یافته این داده‌ها اطلاعات مفیدی در مورد ماهیت بیماری ارائه خواهد داد.



الگوریتم RSA



الگوریتم RSA ، در رمزگذاری داده‌های اینترنت اشیاء مورد استفاده قرار

می‌گیرد. RSA یک الگوریتم رمزنگاری نامتقارن است. این بدان معناست که از یک کلید عمومی و یک کلید خصوصی (یعنی دو کلید متفاوت و ریاضی مرتبط) استفاده می‌کند. کلید عمومی به صورت عمومی به اشتراک گذاشته می‌شود، اما کلید خصوصی مخفی است و نباید با کسی به اشتراک گذاشته شود. RSA مخفف (Rivest-Shamir-Adleman) است که به نام کسانی که در سال ۱۹۷۷ آن را اختراع کردند، نامگذاری شده است.

(n, e) کلید عمومی را تشکیل می‌دهند.

$d - 1 \equiv e^{-1} \pmod{\phi(n)}$ را طوری محاسبه کنید که:

(با الگوریتم اقلیدسی توسعه یافته)

(n, d) کلید خصوصی را تشکیل می‌دهند.

مرحله دوم) رمزگذاری:

یک متن ساده مثل P را در نظر می‌گیریم که به صورت یک عدد نشان داده

می‌شود، متن رمزی C به صورت مقابل محاسبه می‌شود: $C = P^e \pmod{n}$

مرحله سوم) رمزگشایی:

با استفاده از کلید خصوصی متن را با فرمول زیر بدست می‌آوریم:

$$P = C^d \pmod{n}$$

ایده RSA مبتنی بر این واقعیت است که فاکتورسازی یک عدد صحیح

بزرگ دشوار است؛ بنابراین قدرت رمزگذاری کاملاً به اندازه کلید بستگی

دارد و اگر اندازه کلید را دو یا سه برابر کنیم، قدرت رمزگذاری افزایش

می‌یابد. کلیدهای RSA معمولاً می‌توانند ۱۰۲۴ یا ۲۰۴۸ بیت باشند.

الگوریتم‌های رمزگذاری RSA در بسیاری از برنامه‌های کاربردی اینترنت

اشیا استفاده می‌شود. با این حال، بیشتر در گواهینامه های SSL/TLS ،

رمزگذاری ایمیل و ارزشهای دیجیتال یافت می‌شود.

مثال: یک کاربر (به عنوان مثال مرورگر) کلید عمومی خود را به سرویس‌دهنده می‌فرستد و برخی از داده ها را درخواست می‌کند. سرویس‌دهنده داده‌ها را با استفاده از کلید عمومی مرورگر رمزگذاری می‌کند و داده‌های رمزگذاری شده را ارسال می‌کند. مرورگر این داده‌ها را دریافت کرده و رمزگشایی می‌کند. از آنجایی که نامتقارن است، هیچ‌کس دیگری به‌جز مرورگر نمی‌تواند داده‌ها را رمزگشایی کند. حتی اگر شخص ثالث کلید عمومی مرورگر را داشته باشد.

مرحله اول) تولید کلیدها:

۱ - انتخاب دو عدد اول مانند x, y

۲ - ضرب آن دو $n = x * y$

۳ - محاسبه تابع $\phi(n) = (x-1)(y-1)$ totient.

۴ - یک عدد صحیح e را انتخاب کنید، به طوری که نسبت به $\phi(n)$ اول

باشد و $1 < e < \phi(n)$

مهارت‌های مورد نیاز یک مهندس اینترنت اشیا

- برخورداری از دانش فنی عمیق، در ارتباط با کار با کنترلرهای ۳۲ بیتی ARM cortex M۳/M۴
- تجربه کار با پروتکل‌ها و چهارچوب‌های مهم بی‌سیم نظیر ZigBee ، Thread ، یا BLE-Mesh
- دانش تخصصی و تبحر در نوشتن منطق تجاری در ارتباط با تجهیزات توکار با استفاده از روش آزمون محور در پلتفرم‌های RTOS و غیر RTOS
- مهارت کافی در ارتباط با پروتکل‌های ارتباطی نظیر 12C ، SPI ، 1-wire ، UART ، MODBUS ، DALI.
- توانایی حل مسئله و تفکر تحلیلی و تعامل با سایرین. چگونه می‌توان یک مهندس اینترنت اشیا شد؟
 ۱. اصول و مبانی حسگرها را درک کنید.
 ۲. به رابط کاربری توجه کنید.
 ۳. در کار با پایتون و جاوا اسکریپت تجربه کسب کنید.
 ۴. با رزبری پای کار کنید.
 ۵. روی یادگیری مفاهیم بنیادین سیستم عامل‌های موبایل سرمایه‌گذاری کنید.

مهندس اینترنت اشیا از جمله مشاغل این صنعت است که نیازمند مهارت‌های مرتبط با یکدیگر است. مهندسان اینترنت اشیا، افراد متخصصی هستند که می‌توانند با ترکیب سه مولفه اصلی داده، فناوری و تحقیق؛ دستگاه‌ها و سیستم‌های اینترنت اشیا را توسعه داده و بر آن‌ها نظارت و مدیریت کنند. از آنجایی که مهندس اینترنت اشیا، یک متخصص همه فن حریف است، این افراد باید در زمینه‌های مختلف از جمله امنیت، شبکه‌سازی، مهندسی سیستم، برنامه نویسی ابر و برنامه نویسی سخت‌افزاری مهارت کافی داشته باشند.

مهارت‌هایی که یک مهندس اینترنت اشیا باید داشته باشد:

- تجربه برنامه نویسی عملی با زبان‌های برنامه‌نویسی پایتون، Embedded C و Embedded C++.
- درک عمیق از سیستم عامل لینوکس با هدف مدیریت و کنترل دستگاه‌ها و نظارت بر فرآیندهای مورد نیاز.
- آشنایی کامل با معماری نرم‌افزار توکار (embedded)، تجربه کار با نمودارهای کاربردی در UML، نمودار کلاس‌ها (classes diagram) و نمودار توالی (sequence diagram) برای به تصویر کشیدن دقیق معماری سیستمی که قرار است پیاده‌سازی شود.

درس و ابزارها

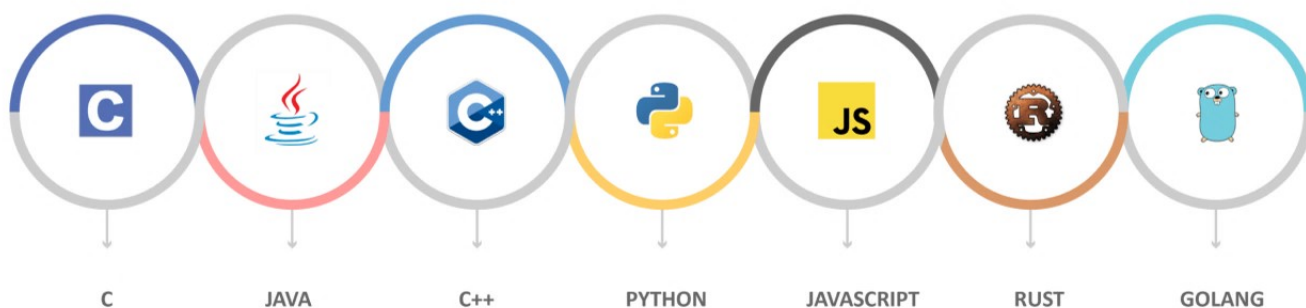
در اینترنت اشیا

• درس:

در حال حاضر این درس در دانشگاه‌های بزرگ جهان در دوره کارشناسی و کارشناسی‌ارشد تدریس می‌شود. در سطح مقدماتی و شروع، درس سخت افزار، طراحی و توسعه نرم افزار، شبکه، هوش مصنوعی، امنیت و داده کاوی مورد نیاز هستند و در ادامه در مراحل بعدی، تمرکز عمیق‌تری بر پروژه‌های مبتنی بر اینترنت اشیا خواهد شد، همچنین با چالش‌ها و قوانین امنیت و آنالیز داده‌های اینترنت اشیا آشنا می‌شوید. از جمله اهداف این رشته آشنایی دانشجویان با مفهوم اینترنت اشیا و کاربردهای آن، آشنایی با معماری چندلایه‌ای اینترنت اشیا و فناوری‌های مرتبط با آن و کسب مهارت در طراحی، پیاده‌سازی و اشکال‌زدایی یک سیستم مبتنی بر اینترنت اشیا ذکر شده است.

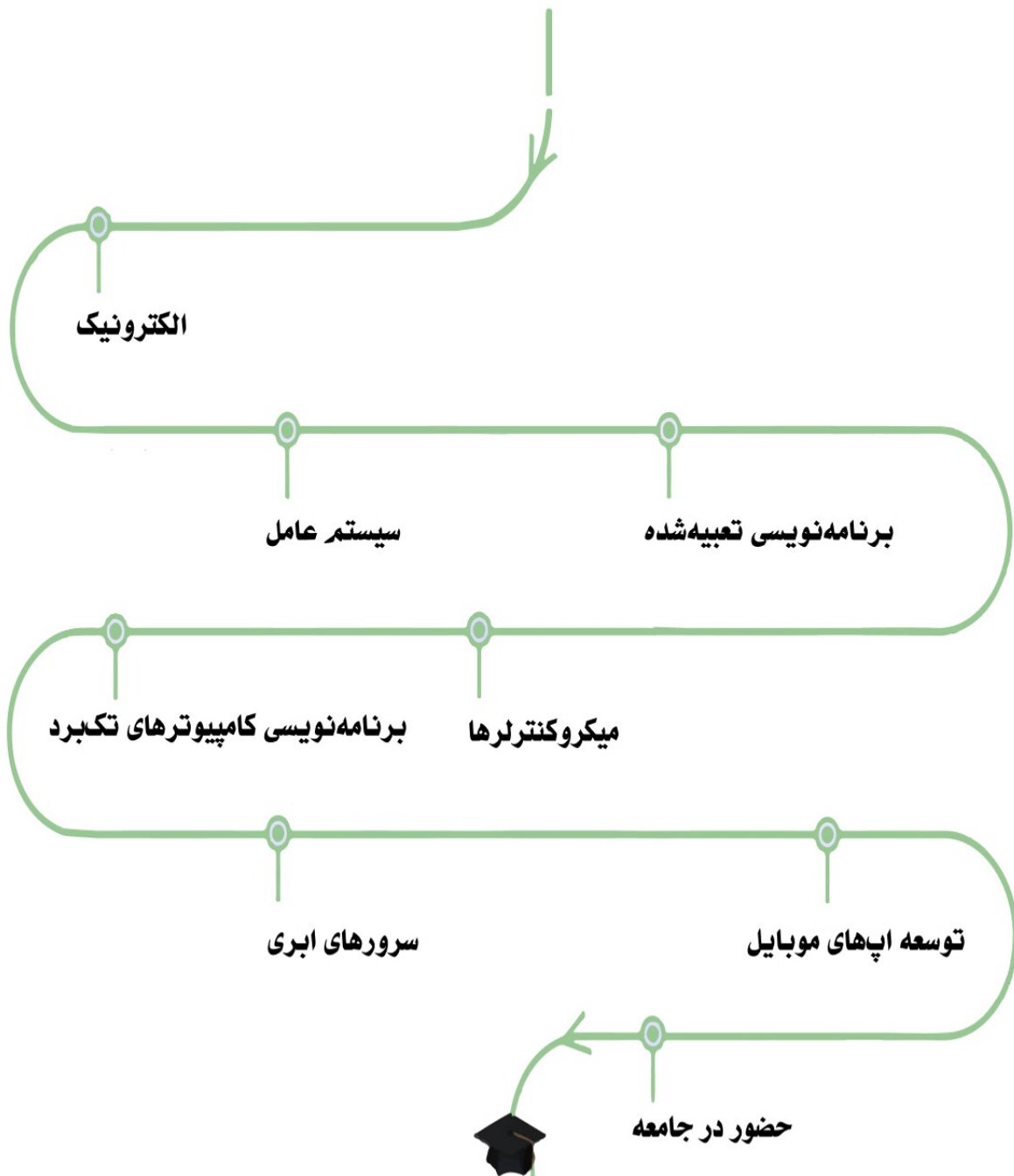
• ابزار:

در رشته اینترنت اشیا JAVA محبوب‌ترین و به نوعی مهم‌ترین زبان برنامه نویسی محسوب می‌شود. یکی از ویژگی‌های برجسته JAVA، مفهوم Write Once, Run Anywhere می‌باشد؛ به این معنی است که کد جاوا همگردانی یا compile شده می‌تواند بر روی هر بستری که از زبان JAVA پشتیبانی می‌کند بدون همگردانی یا ترجمه مجدد اجرا شود. همچنین شی‌گرایی این زبان امکان توسعه برنامه‌ها در هر دو گروه cloud و edge را می‌دهد. بعد از آن زبان‌های C و Python و Javascript از زبان‌های مهم این رشته محسوب می‌شود. Johnny-Five و Cylon.js دو کتابخانه معروف مبتنی بر جاوا اسکریپت در این حوزه‌اند.



نقشه یلگیری مباحث

اینترنت اشیا



معرفی کتاب و دوره

در اینترنت اشیا

کتابهای اینترنت اشیا:

1. IoT and Edge Computing for Architects: Implementing edge and IoT systems from sensors to clouds with communication systems, analytics, and security, by Perry Lea

صنایع برای بهبود هزینه‌های عملیاتی، عمر محصول و رفاه مردم از فناوری‌های IoT استقبال می‌کنند. اگر می‌خواهید از طیف فناوری‌های مورد نیاز برای ساختن یک سیستم اینترنت اشیا موفق، چه یک دستگاه واحد یا میلیون‌ها دستگاه اینترنت اشیا، عبور کنید، به یک راهنمای معماری نیاز است.

این کتاب، طیف گسترده‌ای از راه حل‌های اینترنت اشیا، از حسگرهای اینترنت اشیا گرفته تا ابر را در بر می‌گیرد. برای مثال بخش تجزیه و تحلیل را با استفاده از تحلیل‌های آماری، پردازش رویدادهای پیچیده و مدل‌های یادگیری عمیق تکمیل می‌کند. سپس این کتاب با ارائه دیدگاهی جامع از امنیت اینترنت اشیا، رمزنگاری و امنیت پوسته، علاوه بر محیط‌ها و بلاک‌چین‌های نرم‌افزاری، به پایان می‌رسد.

2. Hands-On Industrial Internet of Things: Create a powerful Industrial IoT infrastructure using Industry 4.0 by Giacomo Veneri and Antonio Capasso

ما در عصری زندگی می‌کنیم که از اتوماسیون پیشرفته برای دستیابی به نتایج دقیق استفاده می‌شود. برای راه اندازی یک محیط اتوماسیون، ابتدا باید شبکه‌ای را پیکربندی کنید که در هر کجا و با هر دستگاهی قابل دسترسی باشد. این کتاب راهنمای عملی است که به شما کمک می‌کند فناوری‌ها و موارد استفاده از اینترنت اشیا صنعتی را کشف کنید.

دوره‌ها:

۱. مقدمه‌ای بر برنامه نویسی تخصص اینترنت اشیا (IoT):

[An Introduction to Programming the Internet of Things \(IoT\) Specialization](#)

این دوره سیستم‌های تعبیه‌شده، با Raspberry Pi و محیط Arduino را برای ساخت دستگاه‌هایی که می‌توانند دنیای فیزیکی را کنترل کنند، پوشش می‌دهد. در پروژه نهایی Capstone، مهارت‌هایی را که با طراحی، ساخت و آزمایش یک سیستم تعبیه‌شده مبتنی بر میکروکنترلر آموخته‌اید، به کار می‌گیرید و یک پروژه نهایی منحصر به فرد را برای نمایش به کارفرمایان آینده تولید می‌کنید.

حتی کد بد هم می‌تواند کار کند.

- خلاصه نکات چهار فصل نخست کتاب کد تمیز -

بخش اول: عمو باب و کتاب کد تمیز

رابرت سسیل مارتین (متولد ۵ دسامبر ۱۹۵۲) معروف به عمو باب، یک مهندس نرم‌افزار، مربی و نویسنده پر فروش آمریکایی است. مارتین کتاب‌ها و مقالات بسیاری در مجلات مختلف نوشته است که یکی از کاربردی‌ترین و مشهورترین آن‌ها کتاب «کد تمیز» است. کتاب کد تمیز مخصوص چه کسانی است؟

در فصل‌های ابتدایی کتاب کد تمیز در مورد مباحث برنامه‌نویسی پایه مانند درست به‌کاربردن توابع، کامنت‌توضیح‌نویسی صحیح، به‌کاربردن صحیح اشیا و درست نوشتن کلاس‌ها صحبت می‌کند که برای افرادی که تازه به برنامه‌نویسی روی آوردند، بسیار کاربردی و مناسب است. در فصل‌های پیشرفته‌تر نیز مباحث ساختار داده‌ها، مدیریت خطا، نحوه به‌کاربردن اصولی تست‌های واحد و نکات مهم در استفاده از هم‌روندی بررسی می‌شوند.

بخش دوم: کد بد

حتی کد بد هم می‌تواند کار کند. اما اگر کد تمیز نباشد، می‌تواند یک شرکت را به زانو درآورد. هر سال، ساعات بی‌شمار و منابع قابل توجهی به‌خاطر کد ضعیف نوشته‌شده از بین می‌روند.

اگر خطاهای کد به مرور زمان رفع نشود و از کد نگهداری نشود، حجم بزرگی از به‌هم‌ریختگی در کد به‌وجود می‌آید و کد به مرور زمان بدتر و بدتر می‌شود، تا جایی که دیگر نمی‌توان کد را مدیریت کرد و همان کد نامناسب باعث ورشکستگی شرکت می‌شود.

فرض کنید یک کد بد جلوی شما قرار گرفته که از پیشرفت شما جلوگیری می‌کند. بهتر است هرچه سریع‌تر برای تمیز کردن کدتان اقدام کنید و شاید به خاطر رسیدن به کارهای دیگران با خودتان بگویید بهتر است این کار را سرهمبندی کنید و بعداً در وقت مناسب آن را تمیز کنید اما طبق قانون LeBlance، بعداً برابری با هیچ‌وقت.

بخش سوم: کد تمیز چیست؟

احتمالاً به اندازه تمام برنامه‌نویسان جهان برای این مفهوم تعریف وجود دارد اما در این بخش، ما برای شما تعریف یک فرد بزرگ در این عرصه را می‌آوریم.

پدرخوانده استراتژی اکلپ «Big Dave Thomas» می‌گوید:

«کد تمیز قابلیت «خوانا بودن» و «بهبودبخشی» توسط توسعه‌دهنده دیگری به غیر از توسعه‌دهنده اصلی کد را دارد، به این منظور کد تست‌های پذیرش و تست‌های واحد دارد، نام‌های بامعنی دارد، یک راه به جای چند راه برای انجام یک کار فراهم می‌کند و حداقل وابستگی را شامل می‌شود. کد باید توضیح مفید داشته باشد، چون بسته به زبان، تمام اطلاعات لازم توسط خود کد به تنهایی قابل بیان نیستند.»

از این دو پاراگراف نتیجه می‌شود که زدن کد تمیز یک هنر است. هم‌چنین نوشتن کد خوب کافی نیست و باید از کد در طول زمان نگهداری کرد تا با گذشت زمان کد خوب به کد بد تبدیل نشود.



بخش چهارم: نام‌گذاری

- از نام‌های آگاهانه و بدون ابهام و قابل تلفظ استفاده کنید:

```

1
2  × int l; //length of array
3  ✓ int arrayLength;
4
5  × int genymdhms;
6  ✓ int generationTimestamp;
7

```

- از نام‌هایی که تفاوت کمی با هم دارند استفاده نکنید:

```

8
9  int ControllerFor
10 × EfficientHandingOfString;
11 int ControllerFor
12 EfficientStorageOfString;
13

```

- اگر در یک Method از list استفاده نکرده‌اید، این کلمه را در نام Method به کار نبرید:

```

14
15 × void accountList() {...}
16 ✓ void accountGroup() {...}
17 ✓ void accountContainer() {...}
18

```

- برای نام‌گذاری از اضافه کردن شماره سری به کلمات دیگر اجتناب کنید:

```

19
20 × int a1, a2, ..., aN;
21

```

- برای نام‌گذاری از کلمات متفاوت که معنای یکسان دارند نیز استفاده نکنید:

```

22
23 × int ProductInfo;
24 int ProductData;
25

```

- از نام‌های قابل جست‌وجو استفاده کنید. پیدا کردن نام‌هایی که تک‌حرفی‌اند یا ثابت‌های عددی کار مشکلی است:

```

26
27 × int a;
28 ✓ int MaxNumber;
29

```

- نام یک کلاس و شیء نباید فعل باشد و باید نام‌های اسمی مانند Customer, Account, Address استفاده شود.
- نام متدها باید یک فعل یا اصطلاح فعلی مانند deletePage, saveData و ... باشد. و یا نام‌هایی مثل accessors بر اساس ارزشی که دارند پشت آن‌ها از پیشوندهایی مثل is, set, get استفاده شود.
- نامها را دقیق انتخاب و سادگی در نام‌گذاری را رعایت کنید و از شوخ‌طبعی در انتخاب نام پرهیز کنید.

```

30
31 × kill();
32 ✓ whack();
33

```

بخش پنجم: توابع

- بهرتر است از تابع‌هایی استفاده کنید که کوچک باشند و از تکرار چند باره‌ی یک کار در تابع خودداری کنید. هرچه تعداد خط‌های یک تابع کمتر باشد، آن تابع بهتر است و همچنین توابع باید حداکثر ۲۰ خط داشته باشند.
- از دیگر مواردی که باید در نوشتن توابع به آن توجه شود، می‌توان به استفاده از بلوک‌ها و برجسته کردن اشاره کرد.
- در تابع باید یک کار را انجام دهید و آن کار را به خوبی انجام دهید. یکی از راه‌های شناختن این که تابع یک کار را انجام می‌دهد، این است که شما نتوانید تابع دیگری با نام دیگر از درون آن بیرون آورید.
- از نام‌های توصیفی استفاده کنید. این نام باید کاری را که تابع انجام می‌دهد، توصیف کند. نام‌های طولانی که توابع را توصیف می‌کنند، بهتر از نام‌های کوتاه مبهم هستند.
- در نوشتن توابع بهتر است تعداد پارامترها کم باشد.

بخش ششم: کامنت‌ها

«کد بد را کامنت نکنید، بلکه آن را دوباره بنویسید.»

“Brian w.Kernighan” and “P.J.Plaugher”

- کامنت عمل ناخوشایندی است که باید انجام شود. اگر زبان برنامه‌نویسی ما به اندازه کافی گویا بود و یا این که ما توانایی استفاده از آن و رساندن منظور خود را داشتیم، دیگر نیازی به نوشتن کامنت نبود.
- کامنت‌ها کد بد را درست نمی‌کنند. به جای صرف وقت برای نوشتن کامنتی که توضیح‌دهنده کدی است که شما ایجاد کرده‌اید، زمان و انرژی را صرف نوشتن کد تمیز کنید.
- کامنت‌های خوب: برخی کامنت‌ها ضروری و مفید هستند و کامنتی خوب

```
// Copyright (C)
// 2003,2004,2005 by Object
// Mentor, Inc. All rights
// reserved.
// Released under the terms
// of the GNU General Public
// License version 2 or later.
```

کامنت‌های آموزنده: گاهی اوقات مفید است که اطلاعات اولیه را به یک کامنت ارایه دهیم:

```
54
55 // Returns an instance of
56 // the Responder being tested.
57 protected abstract Responder
58 responderInstance();
59
```

تا این‌جا خلاصه‌ای از چهار فصل نخست کتاب کد تمیز ارائه شد که ادامه آن را در شماره‌های بعدی بررسی خواهیم کرد.

است که راهی برای نگارش آن پیدا نکنید!

منظورتان را درون کد توضیح دهید. در بسیاری از مواقع چیزی را که می‌خواهید در کامنت بگویید می‌توانید با ایجاد یک تابع بیان کنید:

```
34
35 ✘ // Check to see if the
36 // employee is eligible for
37 // full benefits
38 if ((employee.flags &
39 HOURLY_FLAG) && (employee.age
40 > 65))
41
42 ✔ if (employee.isEligibleFor
43 FullBenefits())
44
```

کامنت‌های قانونی: گاهی اوقات استانداردهای برنامه‌نویسی شرکتی، ما را مجبور می‌کند تا کامنتی خاص به دلایل قانونی بنویسیم. به طور مثال حق نشر و اظهارات نویسنده که در ابتدای هر فایل منع قرار دارد.



«اوگستا آدا بیرون لاولیس» دختر «لرد بیرون» شاعر معروف انگلیسی، مادر برنامه‌نویسی عالم کامپیوتر است. «آدا» در ۱۰ دسامبر سال ۱۸۱۵ در لندن، به دنیا آمد. مادر آدا، تنها شخص در زندگی آدا به شمار میرفت و نقش موثری در کشف استعداد او در ریاضیات داشت.

شهرت آدا بیشتر به واسطه‌ی نوشتن توضیحاتی در مورد ماشین محاسباتی مکانیک «چارلز بابیج» است. آدا نخستین الگوریتم‌نویس کامپیوتری تاریخ محسوب می‌شود.

وی در سال ۱۸۳۵ با «ویلیام کینگ» ازدواج کرد؛ ازدواجی که حاصل آن سه فرزند بود. وی در سن ۳۶ سالگی بخاطر ابتلا به سرطان رحم درگذشت و در کلیسای «سنت ماری ماگدالین» در جوار پدری که هیچ‌گاه شناخت، به خاک سپرده شد.

از جمله رویدادهایی که به نام وی ثبت شد، می‌توان به زبان برنامه‌نویسی آدا، جایزه‌ای که در جامعه‌ی کامپیوتر انگلیس هر سال اهدا می‌شود و مسابقات سالانه‌ای که بین دانشجویان دختر رشته‌های کامپیوتری برگزار می‌شود؛ اشاره کرد. همچنین، در سال ۱۹۹۷ فیلم *Conceiving Ada* با محوریت داستانی «آدا لاولیس» ساخته شد و کتاب «ماشین تفریق» به قلم «ویلیام گیسون» نیز یک رمان تاریخی حول زندگی وی است.