تعهدنامه

اینجانب محدثه دلاوریان دانشجوی دوره کارشناسی ارشد رشته مهندسی کامپیوتر- هوش مصنوعی دانشکده مهندسی دانشگاه فردوسی مشهد نویسنده پایان‌نامه استنباط تقریبی بر روی شبکه بیزی ترکیبی تحت راهنمایی دکتر محمود نقیب زاده متعهد می‌شوم:

- تحقیقات در این پایان‌نامه توسط اینجانب انجام شده و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در پایان‌نامه تاکنون توسط خود و یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه فردوسی مشهد می باشد و مقالات مستخرج با نام "دانشگاه فردوسی مشهد" و یا "Ferdowsi University of Mashhad" به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان‌نامه تاثیرگذار بوده‌اند در مقالات مستخرج از رساله رعایت شده است.
- در کلیه مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است، اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است.

تاریخ ۱۳۹۱/۶/۳۰

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده) متعلق به دانشگاه فردوسی مشهد می‌باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در پایان‌نامه بدون ذکر مرجع مجاز نمی‌باشد.

دانشکده مهندسی

گروه کامپیوتر–گرایش هوش مصنوعی

پایان نامه کارشناسی ارشد

استنباط تقریبی بر روی شبکه بیزی ترکیبی

نویسنده

محدثه دلاوریان

استاد راهنما

جناب آقای پروفسور محمود نقیب زاده

استاد مشاور

جناب آقای دکتر مهدی عمادی

شهریور ۱۳۹۱

دانشگاه فردوسی مشهد
دانشکده مهندسی - گروه مهندسی کامپیوتر

فرم شماره (۷)

صورتجلسه دفاع از پایان‌نامه کارشناسی ارشد

جلسه دفاع از پایان‌نامه کارشناسی ارشد گرایش نرم افزار رشته کامپیوتر در حضور اساتید راهنمای پایان‌نامه، مشاور، مدعوین و نماینده تحصیلات تکمیلی گروه کامپیوتر برگزار شد.
خانم/آقای محدثه دلاوریان پایان‌نامه خود تحت عنوان **استباط تقریبی بروی شبکه بیزی ترکیبی** را ارائه نمودند. ارزیابی اعضای هیات داوران از این پایان‌نامه ۱۸٫۸ به حروف .....

استاد راهنمای پایان‌نامه: دکتر نقیب زاده

استاد مشاور/مدعو: دکتر عمادی

استاد مدعو: دکتر روحانی

نماینده تحصیلات تکمیلی گروه و داور: دکتر منصفی

توضیحات:

خالصانه‌ترین ستایش‌ها در مقام همت والا یتان واسعترین موهبت‌ها را می ستاید. با تقدیم به پدر و مادرم و به پاس زحمات بی دریغشان که شوق زیستن را در طریق بودن به من آموختند

سپاس و قدردانی

از استاد با کمالات و شایسته؛ جناب آقای *پروفسور محمود نقیب زاده* که در کمال سعه صدر، با حسن خلق، از هیچ کمکی در این عرصه بر من دریغ ننمودند و زحمت راهنمایی این پایان نامه را بر عهده گرفتند؛

از استاد صبور، جناب آقای *دکتر مهدی عمادی*، که زحمت مشاوره این رساله را متقبل شدند و راهنمایی‌های ایشان در انجام این پایان نامه بسیار ارزشمند بوده است؛

کمال تشکر و قدردانی را دارم.

باشد که این خردترین، بخشی از زحمات آنان را سپاس گوید.

چکیده

یکی از مدل‌های قدرتمند استنباط، شبکه‌های بیزی هستند. هدف انجام استنباط احتمالی محاسبه احتمال پسین برای متغیرهای تصادفی شبکه می‌باشد. در این پایان‌نامه، تمرکز بر روی این موضوع می‌باشد. در بسیاری از کاربردها، برای مدل کردن مساله متغیرهای تصادفی گسسته و پیوسته به کار می‌رود. شبکه‌ی بیزی مورد مطالعه، در این تحقیق شبکه بیزی ترکیبی است که شامل متغیرهای تصادفی گسسته و پیوسته به صورت همزمان می‌باشد. یکی از چالش‌ها برای استنباط در این شبکه‌ها، متفاوت بودن توابع توزیع احتمال برای انواع متغیرهای تصادفی در آن است که منجر به پیچیدگی بیشتر محاسبات می‌شود. بنابراین اغلب شرایطی برای محدود شدن شبکه و در نتیجه ساده‌تر شدن محاسبات آن در نظر می‌گیرند. در این پایان نامه با در نظر گرفتن این مساله، الگوریتم استنباطی بر روی شبکه‌ی ترکیبی ارائه خواهیم کرد، که بر روی شبکه‌ی ترکیبی استنباط را انجام می‌دهد که برخی از محدودیت‌هایی که در نظر گرفته می‌شود را نداشته باشد. به این هدف به وسیله هم‌شکل سازی شبکه دست می‌یابیم. برای این منظور تابع توزیع هم‌شکلی برای هر دو نوع متغیر تصادفی ارائه شده است. محاسبات الگوریتم را بر اساس تابع توزیع هم‌شکل، برای الگوریتم استنباط که بر پایه‌ی یکی از الگوریتم‌های استنباط تقریبی که انتشار باور Pearl نام دارد، انجام داده‌ایم. برای بررسی نتایج، دقت در تخمین احتمال پسین پس از استنباط و زمان اجرا را برای ارزیابی استفاده می‌کنیم. برای بررسی صحت الگوریتم پیشنهادی آن را بر روی شبکه‌های بیزی گسسته و پیوسته اجرا کرده و نتایج را با الگوریتم‌های دقیق مقایسه کرده و برای شبکه‌های ترکیبی ساده‌تر با الگوریتم‌های دقیق و برای سایر شبکه‌های ترکیبی با روش‌های تقریبی مقایسه کرده‌ایم. نتایج نشان می‌دهند در حالی که شبکه‌ی ترکیبی مورد نظر دارای محدودیت کمتری باشد، با توجه به زمان اجرای الگوریتم پیشنهادی نسبت به روش‌های نمونه برداری به دقت خوبی دست می‌یابیم.

کلمات کلیدی: *شبکه بیزی ترکیبی، استنباط تقریبی، الگوریتم پیام رسانی Pearl، الگوریتم انتشار باور، تابع توزیع شرطی هم‌شکل، الگوریتم پیام رسانی هم‌شکل، تابع توزیع گوسی*

فهرست مطالب

فصل ۱: مقدمه ۵

۱-۱- اهداف ۸

۱-۲- احتمال شرطی و قضیه بیز ۱۰

۱-۳- شبکه ی بیزی ۱۲

۱-۴- انواع شبکه بیزی ۱۶

۱-۵- استنباط ۱۹

فصل ۲: مروری بر کارهای گذشته ۲۲

۲-۱-استنباط دقیق ۲۵

۲-۲- استنباط تقریبی ۲۷

۲-۳- استنباط بر روی شبکه بیزی ترکیبی ۲۸

۲-۴- الگوریتم پیام رسانی Pearl یا انتشار باور ۳۳

۲-۵- تبدیل UT ۳۶

۲-۶- معیار کولبک لیبلر ۳۸

۲-۷- مشکلات و محدودیت ها ۳۸

فصل ۳: الگوریتم پیشنهادی – الگوریتم پیام رسانی هم شکل ۴۱

۳-۱- ایجاد تابع توزیع هم شکل ۴۲

۳-۲- الگوریتم پیام رسانی هم شکل(UnMP) ۵۱

۳-۲-۱- محاسبه پیام $\lambda(X)$ ۵۳

۳-۲-۲- محاسبه پیام $\pi(X)$ ٥٤

۳-۲-۳- محاسبه پیام ارسالی به والدها ٥٦

۳-۲-۴- محاسبه پیام ارسالی به فرزندان ٥٦

فصل ۴: پیاده سازی، نتایج و ارزیابی ۵۸

۴-۱- حالت اول: شبکه ی گسسته ٦٠

۴-۲- حالت دوم: شبکه ی پیوسته ٦٣

٤-۳- حالت سوم: شبکه ی ترکیبی ٦٦

فصل ۵: جمع بندی – نتیجه گیری و توصیه های آتی ۷٦

مراجع ۷۸

پیوست ۱ - محاسبات آماری ۸۲

پیوست ۲- توابع توزیع برخی شبکه ها.....۸۸

فهرست جدول‌ها

جدول ۳-۱- مقادیر احتمال برای متغیر X ۴۵
جدول ۳-۲- مقادیر احتمال ها برای متغیر $U|V$ ۴۶
جدول ۳-۳- مقایسه معیار KL ۵۱
جدول ۴-۱- نتایج بر روی شبکه چند درختی گسسته ۶۲
جدول ۴-۲- نتایج بر روی شبکه *Cancer* ۶۲
جدول ۴-۳- نتایج بر روی شبکه *Asia* ۶۲
جدول ۴-۴- نتایج بر روی شبکه چند درختی پیوسته ۶۳
جدول ۴-۵- نتایج بر روی شبکه *cancer* پیوسته خطی ۶۴
جدول ۴-۶- نتایج بر روی شبکه *cancer* پیوسته غیر خطی ۶۵
جدول ۴-۷- نتایج بر روی شبکه *GHM1* ۶۶
جدول ۴-۸- نتایج بر روی شبکه *GHM2* ۶۸
جدول ۴-۹- نتایج بر روی شبکه *Crop* ۶۹
جدول ۴-۱۰- نتایج بر روی شبکه *Emission* ۷۰
جدول ۴-۱۱- نتایج بر روی شبکه توسعه داده شده *Emission* ۷۲
جدول ۴-۱۲- زمان اجرای الگوریتم ها ۷۳

فهرست شکل‌ها

شکل ۱-۱- شبکه بیزی برای تشخیص وسیله نقلیه. توزیع احتمال محلی مرتبط به هر گره کنار آن نشان داده شده است[*Wei 2007*]. ۱۵

شکل ۱-۲- سمت راست شبکه بیزی چند اتصاله، سمت چپ شبکه بیزی تک اتصاله ۱۶

شکل ۱-۳- شبکه بیزی ترکیبی با سه گره ۱۸

شکل ۲-۱- شبکه بیزی نمونه [*Lan 2009*] ۲۴

شکل ۲-۲- گره x با m والد و n فرزند [*Wei 2007*] ۳۵

شکل ۴-۱- شبکه نمونه چند درختی ۶۱

شکل ۴-۲– شبکه *Cancer* ۶۱

شکل ۴-۳- شبکه *Asia* ۶۱

شکل ۴-۴- شبکه *GHM1* ۶۷

شکل ۴-۵- شبکه *GHM2* ۶۸

شکل ۴-۶- شبکه *Crop* ۶۹

شکل ۴-۷- شبکه *Emission* ۷۱

شکل ۴-۸- شبکه توسعه داده شده *Emission* ۷۲

فصل ۱:

مقدمه

همزمان با رشد سریع قدرت محاسباتی و جمع‌آوری داده در دهه‌های اخیر، روش‌های یادگیری ماشین برای یافتن رابطه‌های پنهان بین داده‌ها به‌کار برده شده‌اند. یکی از مدل‌های قدرتمند احتمالی شبکه‌های بیزی است، که در سال ۱۹۸۰ در حیطه‌ی هوش مصنوعی ظهور کرده‌اند. شبکه بیزی مدل‌سازی ارتباطات پیچیده‌ی بین متغیرهای تصادفی را با ساخت توزیع احتمال توام توسط توزیع احتمال‌های محلی بر عهده دارد. شبکه بیزی یک معماری آشکار برای بیان دانش و استنباط احتمالی دارد. به عنوان یک ابزار مدل‌سازی، شبکه‌های بیزی به ساختار از پیش تعیین شده، تعداد، نوع و رابطه متغیرهای تصادفی محدود نیستند. در سال‌های اخیر توسعه الگوریتم‌های یادگیری ماشین و استنباط، شبکه بیزی را در حیطه حل مسائل احتمالی مشهور کرده است. همچنین قدرتش را در حیطه‌هایی از قبیل تشخیص پزشکی، تشخیص کلاه‌برداری، مهندسی نرم‌افزار، پردازش تصویر و غیره ثابت کرده است.

شبکه‌ی بیزی به خانواده مدل‌های گرافیکی احتمال تعلق دارد. این ساختار گرافیکی برای نمایش دانش در محیط‌های غیر قطعی[1] به کار می‌رود. شبکه‌های بیزی تئوری گراف، تئوری احتمال، علوم کامپیوتر و آمار را با هم ترکیب می‌کند و به یک ساختار شناخته شده مدل‌های گرافیکی با نام گراف جهت‌دار غیر دوری[2] تعلق دارند، که در علم آمار، یادگیری ماشین و هوش مصنوعی شناخته شده هستند. از لحاظ ریاضی پیچیده و به صورت شهودی قابل درک هستند.

هر گره در گراف نشان دهنده یک متغیر تصادفی و یال‌های بین گره‌ها نشان دهنده وابستگی احتمالاتی بین متغیرهای تصادفی مربوطه است. این وابستگی‌های احتمالی در گراف اغلب توسط روش‌های آماری و محاسباتی شناخته شده تخمین زده می‌شوند. [Dar 2008]

[1] uncertain domains
[2] Directed Acyclic Graph (DAG)

شبکه‌های بیزی یکی از محبوب‌ترین و مشهورترین مدل‌ها برای داده‌های چند متغیری است. یادگیری یک شبکه‌ی بیزی از داده‌ها، ساختار احتمالی آن دامنه را آشکار می‌کند و یک ابزار برای پیش‌بینی مشاهدات آینده فراهم می‌آورد.[Sil 2010]

شبکه بیزی مدلی گرافیکی برای بیان ارتباطات احتمالی میان مجموعه‌ای از متغیرها است. در زمینه‌ی تحلیل داده بازنمایی‌های متعددی شامل پایگاه قوانین[۱] درخت‌های تصمیم، شبکه‌های عصبی مصنوعی و همچنین تکنیک‌های متعددی برای تحلیل داده مانند تخمین چگالی[۲]، دسته‌بندی، رگرسیون، و خوشه‌بندی وجود دارد. پس شبکه‌های بیزی و روش بیزی چه پیشنهاد می‌دهند؟[Hec 1995]

- اول، شبکه‌های بیزی به راحتی می‌توانند با داده‌های ناکامل[۳] برخورد کنند، مثلا یک مساله دسته‌بندی و یا رگرسیون را در نظر بگیرید که دارای دو متغیر ورودی به شدت ناهمبسته هستند[۴]. این مسئله برای تکنیک‌های یادگیری با ناظر، در صورتی که تمام ورودی‌ها در هر مورد اندازه‌گیری شده باشند، مشکلی نیست. اما زمانی که یکی از ورودی‌ها مشاهده نشده باشد، اکثر مدل‌ها یک پیش‌بینی نامناسب را تولید می‌کنند. دلیل این است که، آن‌ها وابستگی بین متغیرهای ورودی را لحاظ نمی‌کنند، در حالی که شبکه‌های بیزی یک راه طبیعی برای در نظر گرفتن این وابستگی‌ها پیشنهاد می‌دهند.
- دوم، شبکه‌های بیزی امکان یادگیری در مورد رابطه‌های سببی (علی)[۵] را فراهم می‌آورند. یادگیری در مورد روابط سببی به دو دلیل دارای اهمیت است. فرآیندها هنگامی که سعی داریم درکی از محیط مساله بدست آوریم، مهم هستند. علاوه بر این، داشتن ارتباطات سببی

[1] rule bases
[2] density estimation
[3] incomplete data
[4] anticorrelated
[5] causal relationship

امکان پیش‌بینی در زمان وجود ممانعت[1] را امکان‌پذیر می‌سازند. مثلا یک تحلیل‌گر بازاریابی می‌خواهد بداند که آیا افزایش بودجه تبلیغات جهت بهبود فروش یک محصول توجیه اقتصادی دارد. برای جواب این سوال تحلیل‌گر می‌تواند مشخص کند که آیا تبلیغ یک دلیل برای افزایش فروش است یا خیر. شبکه بیزی کمک می‌کند که به این نوع سوال‌ها جواب داده شود، حتی زمانی که هیچ آزمایشی در مورد اثرات افزایش هزینه موجود نباشد.

- سوم، شبکه‌های بیزی در ارتباط با تکنیک‌های آماری بیزی، ترکیب حیطه‌ی دانش و داده را آسان‌تر می‌سازند.
- چهارم، روش‌های بیزی در رابطه با شبکه‌های بیزی و سایر مدل‌ها یک رویه اصلی و کارا برای جلوگیری از بیش برازش[2] داده فراهم می‌آورند. بنابراین نیازی نیست که برخی از داده‌ها را برای تست نگه داریم و با استفاده از رویه بیزی مدل‌ها می‌توانند تمام داده‌های موجود را برای یادگیری بکار گیرند.

۱-۱- اهداف

در دهه‌ی اخیر شبکه بیزی به عنوان یک ابزار قدرتمند برای بیان عدم قطعیت و استنباط توسط آن نیز مطرح شده است[Jen 2001, Pea 1988]. در بسیاری کاربردها شبکه‌های بیزی دارای متغیرهای تصادفی گسسته و پیوسته هستند، اما به علت نبود الگوریتمی کارا برای استنباط و محاسبات سنگین بر روی این شبکه‌ها در حالت‌های مختلف و عمومی، این حیطه به عنوان یک مساله‌ی باز همچنان مطرح است.

در این پایان نامه بر روی این حیطه از شبکه‌های بیزی تمرکز شده است. هدف در این تحقیق این است که، در یک شبکه بیزی از پیش ساخته شده، با داشتن رخدادی از متغیرهای تصادفی شبکه و یا

[1] intervention
[2] overfitting

بدون هیچ رخدادی احتمال را برای سایر گره‌ها یا به عبارتی متغیرهای تصادفی و یا هر پرس‌وجو در شبکه محاسبه کنیم. به این مرحله کار استنباط بر روی شبکه های بیزی می‌گویند. هدف ما انجام این عمل بر روی نوعی از شبکه‌های بیزی است که ترکیبی خوانده می‌شوند، زیرا شامل متغیرهای تصادفی گسسته و پیوسته می‌باشند.

با توجه به نوع متغیر تصادفی و بیان توزیع احتمال هر کدام از آن‌ها در قالب جدول‌ها[1] و یا توابع توزیع احتمال[2]، استنباط در این شبکه‌ها پیچیده‌تر و دارای محاسبات بیشتری نسبت به سایر شبکه‌ها می‌باشند. بنابراین اغلب شرایطی در نظر می‌گیرند که این شبکه‌ها ساده‌تر شده و عملیات امکان پذیرتر باشند. در این تحقیق متغیرهای تصادفی پیوسته با توابع توزیع گوسی با رابطه خطی و یا غیر خطی را در نظر می‌گیریم. اما بر روی این حالت که یک متغیر تصادفی می‌تواند والد یا فرزند گسسته و یا پیوسته داشته باشد محدودیتی نداریم. ارائه یک الگوریتم استنباط در دسته روش‌های استنباط تقریبی بر پایه الگوریتم پیام رسانی Pearl یا انتشار باور[3] می‌باشد.

برای دستیابی به این هدف، ابتدا برای تمام گره‌ها (متغیرهای تصادفی) تابع توزیع هم‌شکل را پیشنهاد می‌دهیم. سپس بر اساس الگوریتم پایه، پیام‌هایی جهت انتشار احتمال‌ها در بین گره‌های شبکه تعریف می‌شود. در گام بعد مقادیر پیام‌های موجود در الگوریتم محاسبه خواهند شد، که در فصل سوم این محاسبات با جزئیات کامل بیان خواهد شد. سپس برای ارزیابی روش، در سه حالت شبکه کاملا گسسته، کاملا پیوسته و یا ترکیبی روش را اجرا کرده و برای شبکه گسسته و در برخی حالات شبکه پیوسته نتایج با نتایج حاصل از الگوریتم‌های دقیق که بر روی این شبکه‌ها جواب دقیق به دست می‌آورند، مقایسه می‌کنیم. برای شبکه‌های ترکیبی با توپولوژی‌ها و توابع توزیع متفاوت، الگوریتم پیشنهادی با روش‌های تقریبی مقایسه خواهد شد.

---

[1] Probability Table
[2] Probability Distribution Function
[3] Belief Propagation

۱-۲- احتمال شرطی و قضیه بیز

در ادامه به بیان مختصر احتمال شرطی، قانون ضرب و قضیه بیز می‌پردازیم.

۱- احتمال‌های شرطی

در [راس ۱۳۸۶, لاپ ۱۳۸۶] احتمال شرطی به صورت زیر تعریف شده است: "اگر E و F به ترتیب نشان دهنده دو پیشامد باشند. آنگاه احتمال شرطی E به شرط وقوع F به صورت زیر نمایش داده می‌شود: P(E|F). رابطه‌ای کلی برای P(E|F) که برای هر پیشامد E و F معتبر باشد را می‌توان به دست آورد. اگر پیشامد F اتفاق افتد، آنگاه برای اینکه پیشامد E اتفاق افتد لازم است که نتیجه واقعی نقطه‌ای از E و F باشد، یعنی باید در $E \cap F$ قرار داشته باشد. همان‌گونه که می‌دانیم F اتفاق افتاده، در نتیجه فضای نمونه ما به مجموعه F کاهش می‌یابد و بنابراین احتمال اینکه E|F اتفاق افتد برابر است با احتمال وقوع $E \cap F$ نسبت به احتمال وقوع F، که در نتیجه تعریف زیر حاصل می‌شود."

$$P(E|F) = \frac{P(E \cap F)}{P(F)} \qquad (۱-۱)$$

۲- قانون ضرب

در [راس ۱۳۸۶, لاپ ۱۳۸۶] قانون ضرب به صورت زیر تعریف شده است: "اغلب پیشامدها مستقل نیستند. به ویژه، هنگامی که آزمایش‌های تصادفی مربوط به پیشامدهای مورد بحث، وابسته باشند. قانون کلی ضرب را می‌توان به صورت زیر ارائه کرد:

$$P(A \cap B) = P(A)P(B|A)$$

$$P(A \cap B) = P(B)P(A|B) \qquad (۲-۱)$$

توجه داریم که اولین احتمال در سمت راست معادله، یک احتمال غیرشرطی است، در صورتی که دومین احتمال در همان سمت یک احتمال شرطی است. در حالت استقلال پیشامدهای A و B داریم:

$$P(A) = P(A|B) \quad و \quad P(B) = P(B|A) \qquad (۳-۱)$$

قانون ضرب پیشامدها را می‌توان در حالت کلی برای چندین پیشامد تعمیم داد. خواهیم داشت:"

$$P(A_1, A_2, \dots, A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1, A_2) \dots . P(A_n|A_1, A_2, \dots, A_{n-1}) \qquad (٤-۱)$$

۳- قضیه بیز

در [راس ۱۳۸۶, لاپ ۱۳۸۶] قضیه بیز به صورت زیر تعریف شده است: "یک ریاضی‌دان قرن هجدهم به نام توماس بیز[۱]، تعبیر مفاهیم احتمالی را تعمیم داد، بطوری که قضاوت در مورد مقدار احتمال وقوع یک پیشامد، با استفاده از شواهد تجربی انجام می‌شود. روش بیز مبتنی بر مقدار اولیه‌ای از شانس وقوع یک پیشامد است که احتمال پیشین[۲] نام دارد. این عدد ممکن است برحسب نتیجه یک آزمایش تجربی به مقداری کمتر یا بیشتر تجدید نظر شود. البته کم یا زیاد شدن آن بستگی به این دارد که نتیجه یک آزمایش تجربی، پیشامد مورد سوال را تایید یا رد کند. به مقدار احتمال تجدید نظر شده، احتمال پسین[۳] گویند. این روش را می‌توان به وسیله یک فرمول خلاصه و ارائه کرد که به آن قضیه بیز گویند. این قضیه، در تصمیم‌گیری‌های آماری مدرن نقش مهمی بر عهده دارد. در بسیاری از تصمیم‌گیری‌ها، پیشامدهای اصلی، از آزمایش‌های تصادفی تکرار ناپذیر تولید می‌شوند. به نحوی که، قضاوت در مورد ارزیابی احتمالات پیشین، نقش مهمی را ایفا می‌کند. بنابراین، این احتمالات غالبا سلیقه‌ای هستند. به روش‌هایی که احتمالات سلیقه‌ای (ذهنی) را شامل می‌شود، آمار بیزی[۴] می‌گویند.

یک آزمایش تصادفی با دو پیشامد اصلی E و E^c را در نظر بگیرید. احتمالات پیشین مربوط به $P(E^c)$ و P(E) معمولا براساس قضاوت به دست می‌آیند. سعی بر این است که با یک تحقیق تجربی،

[1] Thomas Bayes
[2] Prior Probability
[3] Posterior Probability
[4] Bayesian Statistic

در مورد مقادیر احتمالی بالا تجدید نظری انجام دهیم. فرض شود R نتیجه آزمایش خاصی باشد، احتمالات تجدید نظر شده درباره پیشامدهای اصلی که هدف ما هستند، با شرطی کردن روی R، به صورت احتمالات شرطی $P(E^c|R)$ و $P(E|R)$ به دست می‌آیند.

از آن‌جایی که این احتمالات با تجدید نظر روی احتمالات پیشامدهای اصلی بدست می‌آیند، به آن-ها احتمال پسین می‌گویند. مقادیر احتمالات پسین را می‌توان با استفاده از قضیه بیز به صورت زیر بیان کرد."

$$P(E|R) = \frac{P(E)P(R|E)}{P(R)} = \frac{P(E)P(R|E)}{P(E)P(R|E)+P(E^c)P(R|E^c)} \qquad (۵-۱)$$

قضیه بیز تنها یک فرمول محض نیست بلکه، روشی کارآمد و بسیار مفید در تصمیم‌گیری می‌باشد.

۱-۳- شبکه‌ی بیزی

حال با این مقدمات به تعریف شبکه بیزی می‌پردازیم. یک شبکه بیزی سه گانه‌ای است شامل (V,G,P)، که V مجموعه متغیرهاست (که آن را با X نیز نشان خواهیم داد)، G یک گراف جهت‌دار بدون دور متصل است و یک رابطه یک به یک بین گره‌ها در G و متغیرها در V وجود دارد. P یک مجموعه توزیع‌های احتمال است $P = \{p(v \mid pa(v)) \mid v \in V\}$ که $pa(v)$ نشان‌دهنده مجموعه والدهای V در G است[An 2008]. شبکه‌های بیزی یک بازنمایی و محاسبه توزیع احتمال توام[1] بر روی یک مجموعه متغیرهای تصادفی را امکان پذیر می‌سازند.

شبکه‌ی بیزی تشکیل شده است از: ۱) ساختار شبکه، که یک مجموعه از استقلال‌های شرطی در ارتباط با متغیرهای X در آن وجود دارد. و ۲) یک مجموعه از P، توزیع احتمال محلی مطابق با هر متغیر. این اجزاء با هم توزیع احتمال توام برای X را تعریف می‌کنند [Hec 1995].

[1] Joint probability distribution

ساختار DAG توسط دو مجموعه تعریف می‌شود: مجموعه گره‌ها و مجموعه یال‌های جهت‌دار. گره ها نشان‌دهنده‌ی متغیرهای تصادفی هستند و به صورت مربع و یا دایره که با نام متغیرها برچسب خورده و نمایش داده می‌شوند. یک یال از گره X_i به X_j یک وابستگی آماری بین متغیرهای مربوطه را نشان می‌دهد.

بنابراین یال‌های جهت‌دار نشان می‌دهند که ارزش متعلق به متغیر X_j به ارزش متعلق به متغیر X_i وابسته است و یا به عبارتی X_i بر روی X_j تاثیر می‌گذارد. گره X_i به عنوان والد X_j، X_j به عنوان فرزند X_i شناخته می‌شود. یک توسعه از آن می‌تواند به عنوان مجموعه اولاد[۱]، مجموعه گره‌هایی که با یک مسیر مستقیم از یک گره در دسترس هستند، و گره‌های اجداد[۲]، مجموعه گره‌هایی که از یک گره توسط یک مسیر مستقیم می‌تواند در دسترس باشد، مطرح شود[Ben 2007]. یال‌های گراف این حکم را بیان می‌کنند که یک متغیر به صورت شرطی مستقل از غیر اولادهایش[۳] در گراف است هنگامی که والدینش داده شده باشند.

از نظر ساختار، شبکه‌های بیزی به دو نوع تک اتصاله[۴] و چند اتصاله[۵] تقسیم می‌شوند. در شبکه تک اتصاله که چند درختی[۶] نیز نامیده می‌شود، بیشتر از یک مسیر[۷] بین هر دو گره وجود ندارد. از طرف دیگر در شبکه‌های بیزی چند اتصاله حداقل بین یک جفت از گره‌ها بیشتر از یک مسیر وجود دارد (گراف دارای دور بدون جهت است).

علاوه بر ساختار DAG که اغلب به عنوان قسمت کیفی[۸] مدل در نظر گرفته می‌شود، نیاز است که پارامترهای کمی[۱] مدل را نیز مشخص کنیم. پارامترها به گونه‌ای تعریف می‌شوند که با خاصیت

---

[1] descendants
[2] ancestor
[3] non-descendant
[4] singly-connected
[5] multiply-connected
[6] polytree
[7] path
[8] qualitative