



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)
دانشکده مهندسی کامپیوتر و فناوری اطلاعات

گزارش ارائه درس داده کاوی

الگوریتم بیشینه سازی انتظار (Expectation Maximization)

نگارش

عارف معتمدی

حسین بهشتی فرد

سجاد سرلکی

استاد مربوطه

دکتر احسان ناظر فرد

خرداد ۱۴۰۰

اینجانب عارف معتمدی، حسین بهشتی فرد، سجاد سرلکی مراتب امتنان خود را نسبت به استاد گرانقدر آقای دکتر احسان ناظر فرد که طی تولید ارائه و تدوین این گزارش، همواره ما را یاری نموده‌اند و همچنین تمام دوستان و افرادی که مرا در پروسه این تحقیق و پژوهش یاری کرده‌اند، ابراز می‌دارم.

عارف معتمدی، حسین بهشتی فرد، سجاد سرلکی

خرداد ۱۴۰۰

چکیده

همانطور که می‌دانید یک دسته مهم از الگوریتم‌های هوش مصنوعی، دسته الگوریتم‌های خوشه بندی هستند. خود دسته الگوریتم‌های خوشه بندی به از دیدگاه‌های مختلفی به بخش‌هایی تقسیم می‌شوند.

از یک دیدگاه این الگوریتم‌ها به الگوریتم‌های "خوشه بندی سخت"^۱ و "خوشه بندی نرم"^۲ تقسیم می‌شوند. از دیدگاه دیگر بخشی الگوریتم‌های "خوشه بندی بر مبنای مدل"^۳ است.

امروزه مدل‌های مبتنی بر احتمال^۴ به منظور شناخت ساختار اطلاعات و داده‌ها در علوم مختلف به کار می‌روند. یکی از این کاربردها استفاده در یادگیری ماشین و بحث خوشه بندی است و به طور ریزتر در بحث خوشه بندی بر مبنای مدل – که پیشتر گفتیم – است. این استفاده به طور عملی در الگوریتم EM قابل مشاهده است. در این گزارش به معرفی و بررسی الگوریتم EM می‌پردازیم.

واژه‌های کلیدی:

خوشه‌بندی، بیشینه رخداد، الگوریتم بیشینه‌سازی انتظار

Hard-Clustering^۱

Soft-Clustering^۲

Model Based Clustering^۳

Probabilistic Model^۴

فصل اول مقدمه.....	۱
فصل دوم پیدا کردن بیشینه احتمال رخداد.....	۴
۱-۲- تعریف و مفهوم بیشینه احتمال رخداد.....	۵
۲-۲- پیدا کردن بیشینه رخداد در سیستم ها.....	۶
۱-۲-۲- پیدا کردن بیشینه رخداد در سیستم های با نظارت.....	۶
۱-۲-۲- تفکر برعکس.....	۷
۲-۲-۲- پیدا کردن بیشینه رخداد در سیستم های بی نظارت.....	۸
فصل سوم الگوریتم بیشینه سازی انتظار آشنایی.....	۹
۱-۳- مفهوم و تعریف الگوریتم بیشینه سازی انتظار.....	۱۰
۲-۳- اهمیت الگوریتم بیشینه سازی احتمال و وجه تمایز آن.....	۱۰
۳-۳- الگوریتم بیشینه سازی انتظار.....	۱۲
۳-۳-۱- مقدار دهی اولیه.....	۱۳
۳-۳-۲- پیدا کردن انتظار.....	۱۴
۳-۳-۳- بیشینه کردن انتظار.....	۱۵
۳-۳-۴- ارزیابی همگرایی الگوریتم.....	۱۶
۳-۴- مزایا و معایب.....	۱۷
فصل چهارم کاربرد های EM.....	۱۹
۱-۴- کاربرد در رشته های مختلف.....	۲۰
۲-۴- بررسی مثالی ساده ::.....	۲۱
فصل پنجم جمع بندی و نتیجه گیری و پیشنهادات.....	۲۲
منابع و مراجع.....	۲۴

صفحه

فهرست اشکال و جداول

شکل ۱ - انواع توزیع های آماری	۲
شکل ۱ - توزیع نورمال	۳
شکل ۳ - Maximum Likelihood	۵
شکل ۴ - Maximum Likelihood Estimation	۶
شکل ۵ - داده ها در یادگیری با نظارت	۷
شکل ۶ - پیدا کردن مدل آماری در یادگیری با نظارت	۷
شکل ۷ - داده های بدون برچسب در توزیع های مشخص شده	۸
شکل ۸ - روی هم افتادگی دو توزیع احتمالاتی	۱۱
شکل ۹ - خوشه بندی انجام شده توسط دو الگوریتم k-means و EM بر روی مجموعه داده موش	۱۲
شکل ۱۰ - فلوجارت الگوریتم پیشنهادی سازی انتظار	۱۳
شکل ۱۱ - مجموعه داده های آزمایشی و تجسم پارامتر های مقداری دهی شده اولیه	۱۴
شکل ۱۲ - مجموعه داده های آزمایشی و تجسم داده ها پس از محاسبه احتمال تعلق	۱۵
شکل ۱۳ - مجموعه داده های آزمایشی و تجسم داده ها پس از بروزرسانی پارامتر های الگوریتم	۱۶
شکل ۱۴ - مجموعه داده های آزمایشی و اجرای ۲۰ تکرار از الگوریتم تا همگرایی	۱۷
شکل ۱۵ - رشد احتمال رخداد لگاریتمی به ازای هر تکرار از الگوریتم	۱۷
شکل ۱۶ - خروجی فشرده سازی	۲۱
شکل ۱۷ - مقایسه عملکرد EM و FCM	۲۱

فصل اول

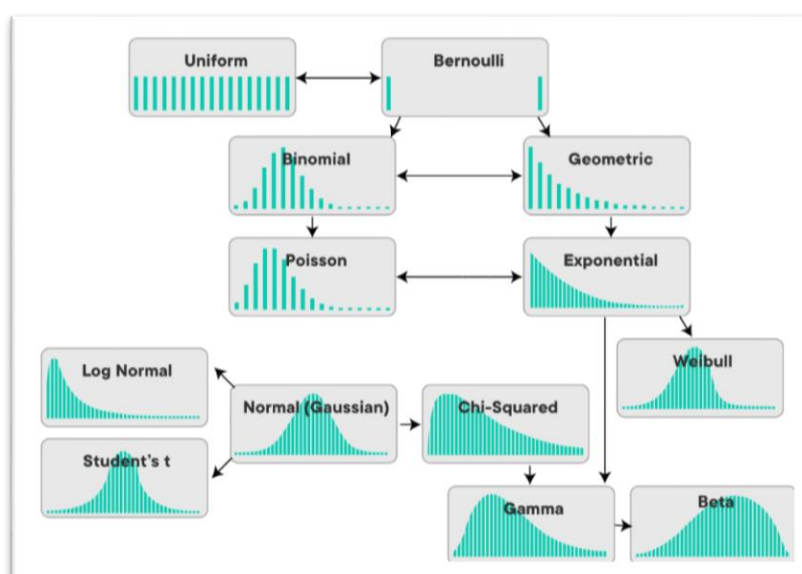
مقدمه

مقدمه

همانطور که می‌دانید برخی قلب علم کار و تحلیل داده را آمار و ریاضیات می‌دانند. در عمل بسیاری از محققین، علم داده را یک مثلث با سه ضلع علوم کامپیوتر، آمار و ریاضیات و علم کسب و کار می‌دانند. در علم داده‌ها یکی از بزرگترین نیازهای ما دانستن مدل آماری داده‌ها است تا بتوان یک بینش تصویری از مجموعه خود داشته باشیم. همینطور که گفتیم اثبات اهمیت علم آمار در مبحث علم داده‌ها بر کسی پنهان نیست.

آندره دانکن ریاضی دان مشهور می‌گوید "شاید بتوانیم با آمار دروغ بگوییم. ولی هرگز نمی‌توانیم بدون آمار راست بگوییم".

در علم آمار ما نیاز داریم تا توزیع^۵ داده‌هایمان را بدانیم. توزیع‌های آماری گوینده روند پخش داده‌ها، نقاط پرازدهام تر، مرکز داده‌ها و ... است. ما در علم ریاضی و آمار توزیع‌های آماری متفاوتی داریم که مهم‌ترین آن، توزیع نورمال^۶ است.



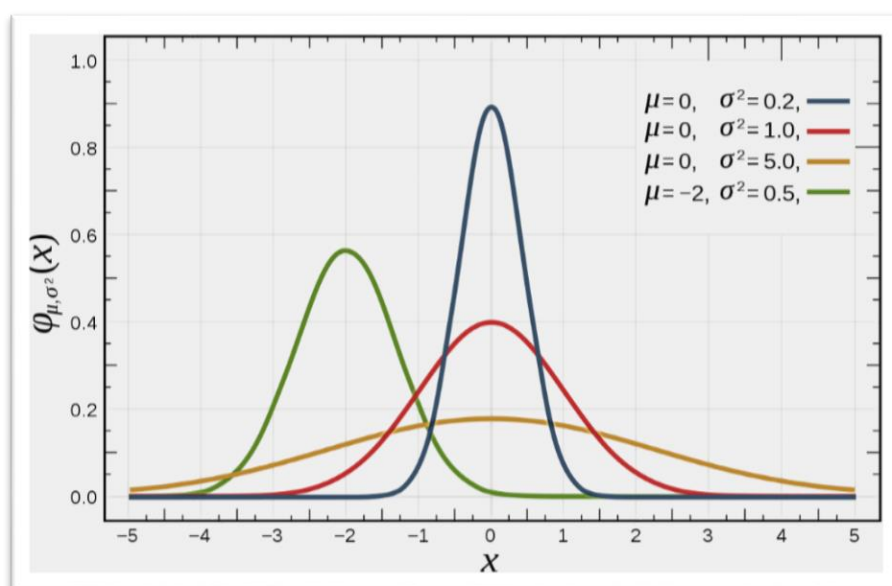
شکل ۲ - انواع توزیع‌های آماری

^۵ Distribution

^۶ Normal distribution

همانطور که در شکل ۱ می‌بینید، توزیع‌های مهمی مانند توزیع یگانه^۷، نمایی^۸، لگاریتمی نمایی^۹ و ... داریم.

همانگونه که گفتیم مهم‌ترین توزیع آماری برای ما، توزیع آماری نورمال است. در این توزیع (شکل ۲) داده‌ها حول یک مرکز به نام میانگین، پخش شده‌اند و هرچه از میانگین^{۱۰} دورتر می‌شویم، به صورت متقارن داده‌ها کمتر و کمتر می‌شوند. پارامتری به نام انحراف معیار^{۱۱} و یا واریانس^{۱۲} (که توان ۲ انحراف معیار است) نشان‌دهنده تراکم این داده‌ها در اطراف میانگین است.



شکل ۳ - توزیع نورمال

در این ارائه تمام تمرکز ما روی توزیعهای نورمال و پیدا کردن واریانس و میانگین مجموعه‌ای از داده است.

uniform ^۷

exponential ^۸

Log exponential ^۹

mean ^{۱۰}

Standard deviation ^{۱۱}

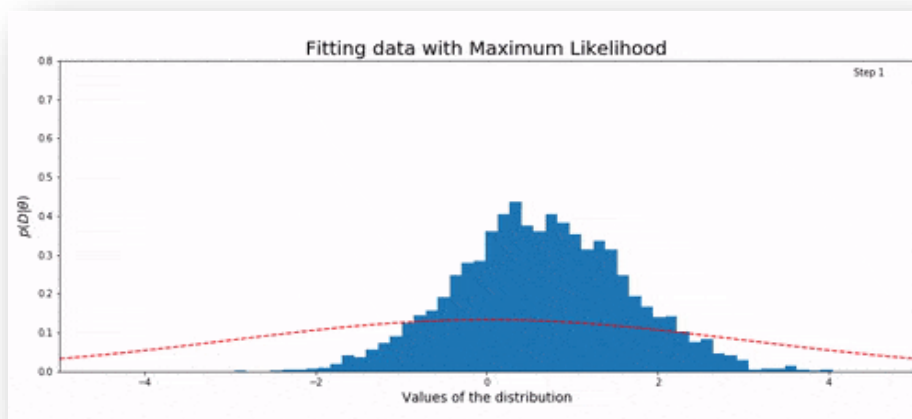
variance ^{۱۲}

فصل دوم

پیدا کردن بیشینه احتمال رخداد

پیدا کردن بیشینه احتمال رخداد^{۱۳}

خوب است ابتدا پیش از ورود به بحث اصلی و باز کردن الگوریتم بیشینه سازی انتظار، با مفهوم بیشینه احتمال رخداد آشنا شویم. در این فصل ما این مفهوم را تعریف می‌کنیم و سپس سعی می‌کنیم در سیستم‌ها این بیشینه را پیدا کنیم. پیدا کردن این مفهوم را در سیستم‌های یادگیری با نظارت و بدون نظارت بررسی می‌کنیم و می‌یابیم که چرا در سیستم‌های بدون نظارت بسیار مشکل داریم و مجبوریم به سراغ الگوریتم‌های پیچیده‌تر مانند بیشینه سازی انتظار برویم.

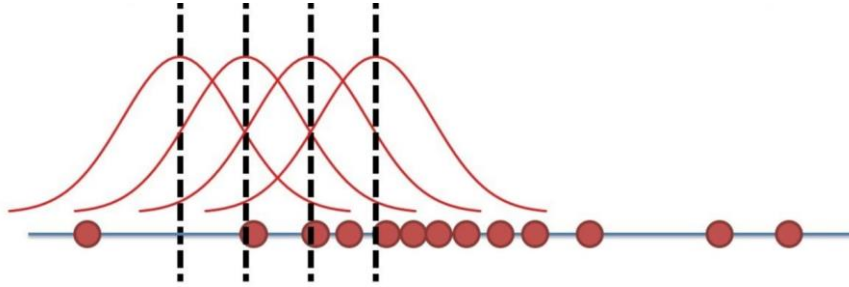


شکل ۴ - Maximum Likelihood

۲-۱- تعریف و مفهوم بیشینه احتمال رخداد

همانطور که از اسم این مفهوم می‌توان درک کرد، احتمال رخداد یک مجموعه داده به مرکز C_i را در نقطه x_i ، احتمال رخداد آن مجموعه در نقطه x_i می‌گویند. اگر بیایم این احتمال را برای تمام نقاط x_i به ازای $x_i \in D$ حساب کنیم و بیشینه آنها را پیدا کنیم، می‌توان گفت احتمال رخداد این مجموعه در نقطه بدست آمده از همه بیشتر است و این نقطه دارای بیشینه احتمال رخداد آن مجموعه داده است.

پیدا کردن بیشینه رخداد^{۱۴} روشی است در آن پارامترهای توزیع داده (مانند میانگین و واریانس) در یک مجموعه حساب می شود به طوری که در یک مدل آماری، داده های مشاهده شده ما احتمال بیشینه ای داشته باشند.



شکل ۵ - Maximum Likelihood Estimation

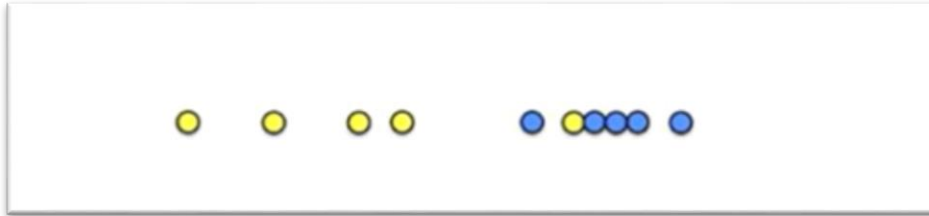
۲-۲- پیدا کردن بیشینه رخداد در سیستم ها

حال می خواهیم در سیستم های متفاوت این روش را حساب کنیم. سیستم های خود را به دو گروه سیستم های دارای یادگیری با نظارت و بدون نظارت تقسیم بندی می کنیم. در سیستم های یادگیری با نظارت، داده ها برچسب دارند و مجموعه، می دانیم هر داده به کدام دسته متعلق است. ولی در سیستم های بدون نظارت، داده های ما برچسب گروه ندارند. در نتیجه کار کمی دشوار تر می شود.

۲-۲-۱- پیدا کردن بیشینه رخداد در سیستم های با نظارت

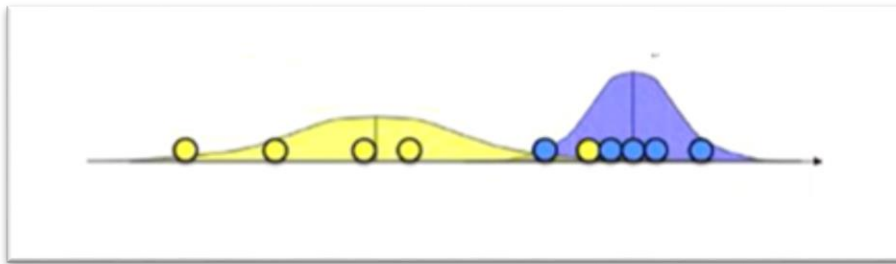
همانطور که گفتیم، در سیستم های یادگیری با نظارت، داده های ما برچسب گروه های خود را دارند. به این حالت که می دانیم هر داده، متعلق به کدام دسته است. حال در مسئله های پیدا کردن بیشینه رخداد در این سیستم ها، ما طریقه دسته بندی داده ها را داریم و تنها باید مدل آماری را بدست آوریم.

به طور مثال به شکل ۵ توجه کنید. ما دو دسته داده زرد و آبی داریم. حال می‌خواهیم نمود آماری هر دسته را پیدا کنیم.



شکل ۶ - داده ها در یادگیری با نظارت

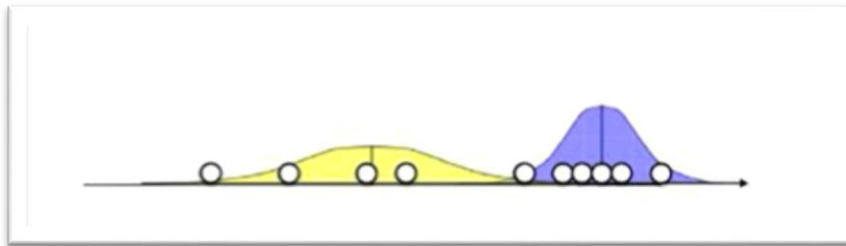
بدیهیست که با این دسته بندی داده ها می‌توان میانگین هر کدام از آنان را پیدا کرد و طبق میانگین و اعضای هر دسته، واریانس را طبق فرمول های گفته شده پیدا کرد. در نتیجه ما در انتها توزیع نورمال مربوط به این داده ها را پیدا می‌کنیم.



شکل ۷ - پیدا کردن مدل آماری در یادگیری با نظارت

۲-۱-۱-۲- تفکر برعکس

قبل رفتن به مبحث بعدی و پیدا کردن این بیشینه در سیستم های بی نظارت، بیاییم برعکس فکر کنیم. ما در اینجا کلاس بندی هر داده را داشتیم و طبق کلاس هر داده، توزیع و مدل آماری را بدست آوردیم. حال اگر مسئله ما برعکس باشد، یعنی توزیع آماری را داشته باشیم و بخواهیم به داده‌هایمان طبق کلاس ها برچسب بزنیم، مسئله ای مثل زیر داریم.



شکل ۸ - داده های بدون برچسب در توزیع های مشخص شده

در این نقطه، برای هر داده، طبق میانگین و واریانس هر کلاس چک می کنیم احتمال رخداد تعلق این داده به هر کلاس چقدر است و در انتها هر داده بنا بر احتمال کسری به هر یک از کلاس ها متعلق می شود.

۲-۲-۲- پیدا کردن بیشینه رخداد در سیستم های بی نظارت

در این مدل یادگیری ها، داده های ما نه برچسب دارند و نه توزیع آماری آن ها پیدا است. در قسمت قبل ما دیدیم که با داشتن یکی از موارد گفته شده، برچسب و یا توزیع آماری، می توان به دیگری رسید. ولی مسئله ای که به وجود می آید این است که در این قسمت ما هیچکدام را نداریم. در نتیجه مسئله ما نمونه جدیدی از مسئله مرغ و یا تخم مرغ است که باید ببینیم کدامیک اول رخ می دهد.

در این مسئله ها ما نیازمند به الگوریتم های پیچیده تر داریم. یکی از این الگوریتم ها EM و یا بیشینه سازی انتظار است که در فصول بعدی به آن می پردازیم.

فصل سوم

الگوریتم بیشینه سازی انتظار

آشنایی با الگوریتم بیشینه‌سازی انتظار^۱

ما تا به اینجای کار با تعدادی از مفاهیم پیش‌نیاز آماری و بخصوص بیشینه رخداد، آشنا شده‌ایم. حال به سراغ هدف اصلی خود که بررسی الگوریتم بیشینه‌سازی انتظار می‌باشد، می‌رویم.

۳-۱- مفهوم و تعریف الگوریتم بیشینه‌سازی انتظار

بطور کلی الگوریتم بیشینه‌سازی انتظار، یک الگوریتم برپایه‌ی تکرار می‌باشد که برای پیدا کردن بیشینه تعداد رخداد^۲ و یا پیدا کردن بیشینه احتمال خلفی^۳، مورد استفاده قرار می‌گیرد. زمانی که به دنبال پیدا کردن الگوها و توزیع‌های احتمالی و یا تفکیک خوشه‌ها هستیم، این الگوریتم می‌تواند بسیار مفید واقع شود.

خوشه‌بندی در این الگوریتم به صورت نرم^۴ می‌باشد و در پایان الگوریتم برای هر داده مجموعه‌ای از احتمالات تعلق^۵ به خوشه‌های متفاوت خواهیم داشت.

۳-۲- اهمیت الگوریتم بیشینه‌سازی احتمال و وجه تمایز آن

امروزه با پیشرفت چشم‌گیر هوش مصنوعی، الگوریتم‌های متنوعی برای خوشه‌بندی و پیدا کردن الگوها پدید و مورد استفاده قرار می‌گیرند، که یکی از مهم‌ترین و متداول‌ترین آن‌ها الگوریتم k-means می‌باشد.

^۱ Expectation maximization

^۲ Maximum likelihood

^۳ Maximum posterior probability

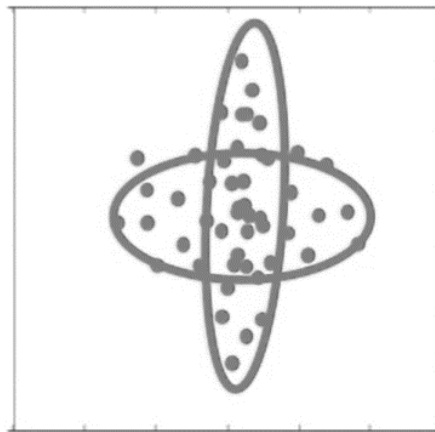
^۴ Soft clustering

^۵ Responsibilities

حال این سوال پیش می‌آید که الگوریتم بیشینه‌سازی انتظار چه تفاوتی با این الگوریتم‌های متداول دارد و آیا برتری‌ای وجود دارد یا خیر؟

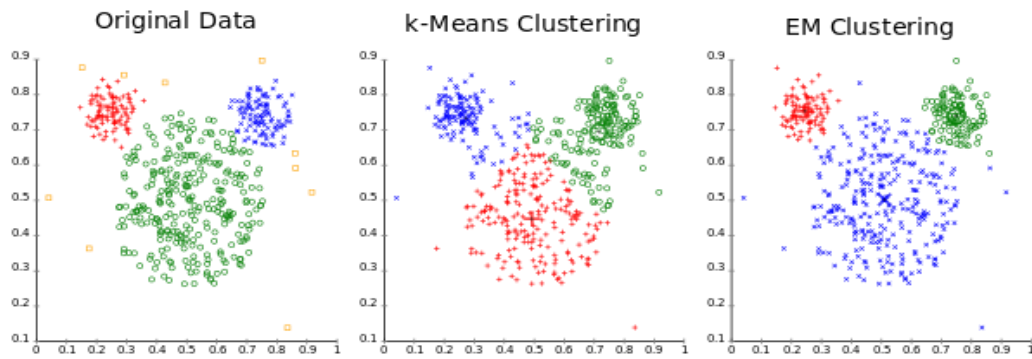
تفاوت اصلی این الگوریتم با الگوریتم‌های خوشه‌بندی دیگر مانند k-means که اغلب از فاصله اقلیدوسی و منهتن استفاده می‌کنند، این است که این الگوریتم از آمار و احتمالات و قانون بیز استفاده می‌کند که باعث می‌شود توزیع‌های احتمالی را به خوبی تشخیص دهد.

فرض کنیم مجموعه داده‌های ما تشکیل شده از تعداد چندین توزیع احتمالی باشد. زمانی که مرکز این توزیع این خوشه‌ها مشترک باشد الگوریتم k-means چگونه خوشه‌بندی خواهد کرد؟ اگر از الگوریتم‌های غیرخطی مانند DBSCAN استفاده کنیم این تعداد توزیع‌ها چه نتیجه‌ای را دربرخواهد داد؟



شکل ۸ - روی هم افتادگی دو توزیع احتمالاتی

همینطور که می‌دانیم بسیاری از این الگوریتم‌هایی که بر اساس فاصله عمل می‌کنند، نمی‌توانند نتیجه خوبی برای این مدل‌ها ارائه دهند. اما الگوریتم بیشینه‌سازی انتظار این توزیع‌ها را با دقت بالایی تشخیص می‌دهد.



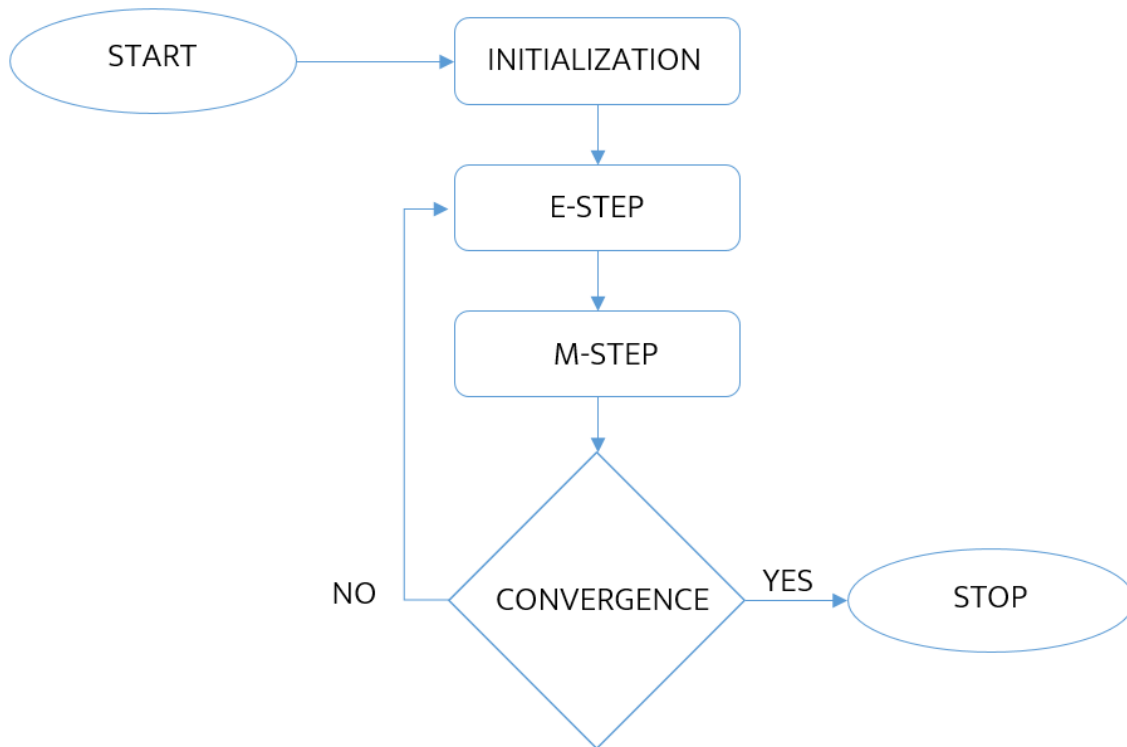
شکل ۹ - خوشه بندی انجام شده توسط دو الگوریتم k-means و EM بر روی مجموعه داده موش

در ادامه با روند این الگوریتم آشنا خواهیم شد.

۳-۳- الگوریتم بیشینه سازی انتظار

الگوریتم بیشینه سازی انتظار سه پارامتر اصلی μ_k که نشان دهنده میانگین، Σ_k نشان دهنده کواریانس و π_k ضریب اختلاط، دارد. بطور کلی این الگوریتم از دو مرحله اصلی محاسبه میزان انتظار و بیشینه کردن این انتظار تشکیل شده که پس از مقداردهی اولیه این دو مرحله تا زمان همگرایی الگوریتم تکرار خواهند شد.

فلوچارت این الگوریتم بصورت زیر می باشد که این مراحل را در ادامه بررسی خواهیم کرد.

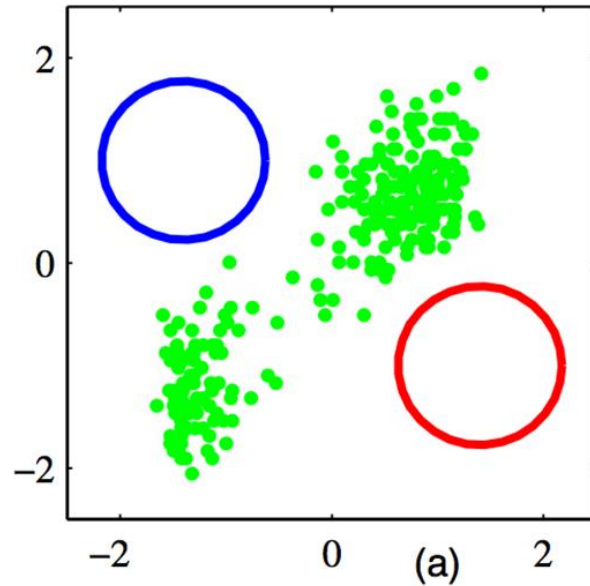


شکل ۱۰ - فلوچارت الگوریتم بیشینه‌سازی انتظار

۱-۳-۳- مقدار دهی اولیه

در ابتدا نیاز است که سه پارامتر μ_k ، \sum_k و π_k را مقدار دهی کنیم.

این مقدار دهی می‌تواند بصورت تصادفی باشد. اما معمولاً از الگوریتم‌های دیگر بخصوص الگوریتم k -means برای مقدار دهی اولیه این پارامترها استفاده می‌شود. به اینصورت که در ابتدا یک بار الگوریتم k -means اجرا می‌شود و پس از پایان الگوریتم و مشخص خوشه‌ها میانگین و توزیع آن‌ها سپس الگوریتم بیشینه‌سازی انتظار را اجرا می‌کنیم.



شکل ۱۱ - مجموعه داده های آزمایشی و تجسم پارامتر های مقداری دهی شده اولیه

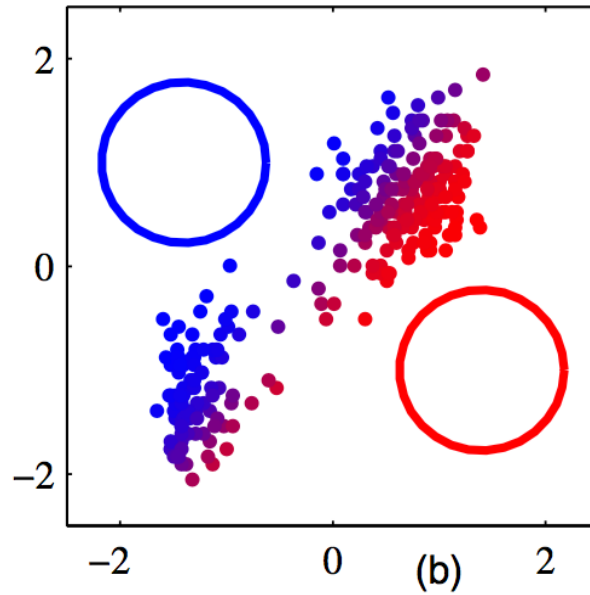
۲-۳-۳- پیدا کردن انتظار

در این مرحله برای هر داده n و خوشه k پارامتر γ را که نشان دهنده ی احتمال تعلق می باشد را بصورت زیر با استفاده از قانون بیز محاسبه می کنیم.

$$\begin{aligned}\gamma_k = p(z = k|\mathbf{x}) &= \frac{p(z = k)p(\mathbf{x}|z = k)}{p(\mathbf{x})} \\ &= \frac{p(z = k)p(\mathbf{x}|z = k)}{\sum_{j=1}^K p(z = j)p(\mathbf{x}|z = j)} \\ &= \frac{\pi_k \mathcal{N}(\mathbf{x}|\mu_k, \Sigma_k)}{\sum_{j=1}^K \pi_j \mathcal{N}(\mathbf{x}|\mu_j, \Sigma_j)}\end{aligned}$$

$$\gamma_k^{(n)} = p(z^{(n)}|\mathbf{x}) = \frac{\pi_k \mathcal{N}(\mathbf{x}^{(n)}|\mu_k, \Sigma_k)}{\sum_{j=1}^K \pi_j \mathcal{N}(\mathbf{x}^{(n)}|\mu_j, \Sigma_j)}$$

در انتهای این مرحله با توجه به پارامتر γ محاسبه شده احتمال تعلق هر داده را نسبت به هر خوشه خواهیم داشت.



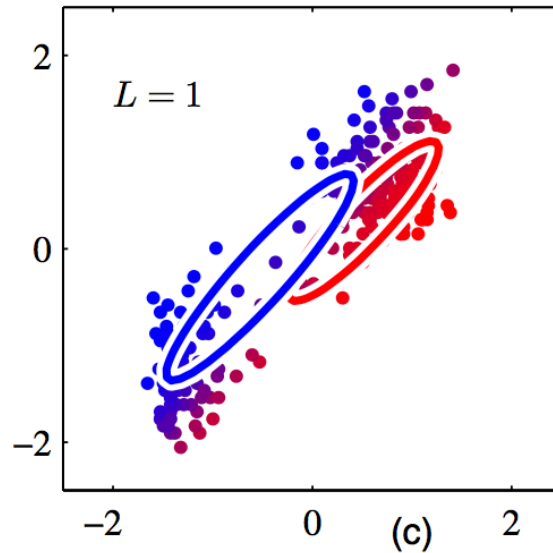
شکل ۱۲ - مجموعه داده های آزمایشی و تجسم داده ها پس از محاسبه احتمال تعلق

۳-۳-۳- بیشینه‌کردن انتظار

در این مرحله براساس احتمال که در مرحله قبل بدست آوردیم سه پارامتر μ_k ، Σ_k و π_k را بصورت زیر بروزرسانی می‌کنیم.

$$\begin{aligned}\mu_k &= \frac{1}{N_k} \sum_{n=1}^N \gamma_k^{(n)} \mathbf{x}^{(n)} \\ \Sigma_k &= \frac{1}{N_k} \sum_{n=1}^N \gamma_k^{(n)} (\mathbf{x}^{(n)} - \mu_k)(\mathbf{x}^{(n)} - \mu_k)^T \\ \pi_k &= \frac{N_k}{N} \quad \text{with} \quad N_k = \sum_{n=1}^N \gamma_k^{(n)}\end{aligned}$$

پس از پایان این مرحله میانگین و کواریانس به سمت داده هایی که تعلق بیشتری دارند، حرکت می‌کنند.



شکل ۱۳ - مجموعه داده‌های آزمایشی و تجسم داده‌ها پس از بروزرسانی پارامترهای الگوریتم

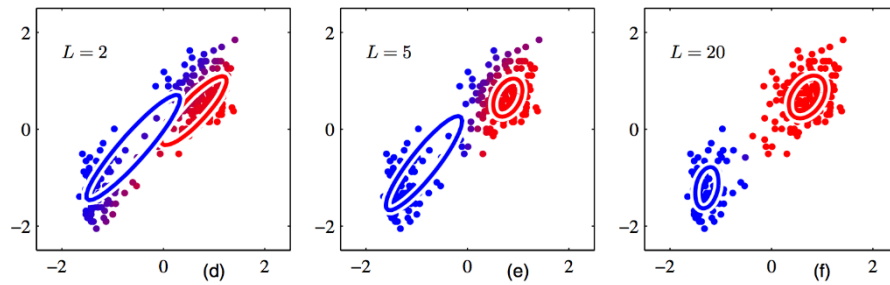
۴-۳-۳- ارزیابی همگرایی الگوریتم

برای ارزیابی و چک کردن همگرا شدن احتمال رخداد لگاریتمی^۱ بصورت زیر استفاده می‌شود.

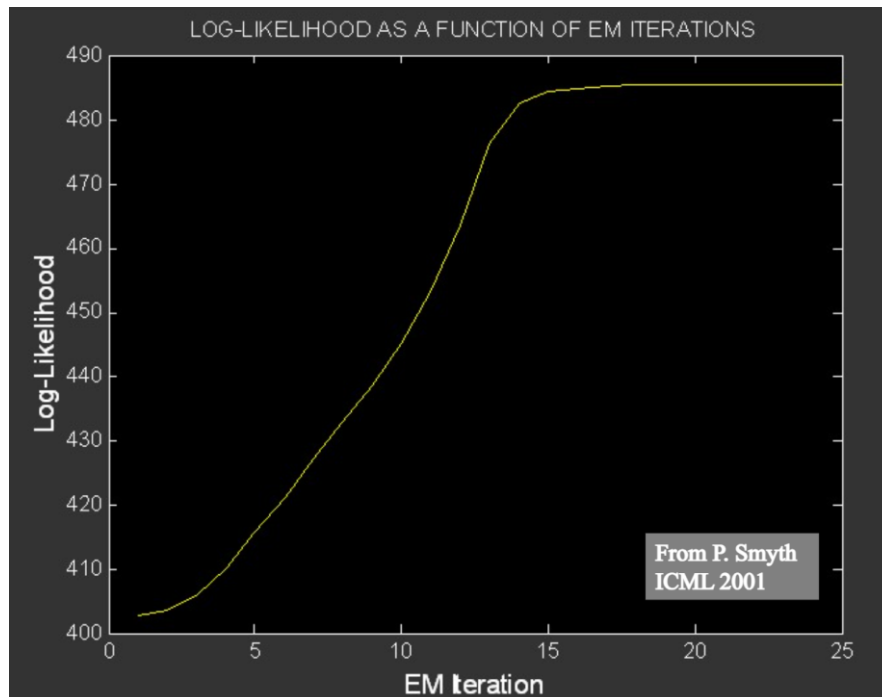
$$\ln p(\mathbf{X}|\pi, \mu, \Sigma) = \sum_{n=1}^N \ln \left(\sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{x}^{(n)}|\mu_k, \Sigma_k) \right)$$

با پیشرفت الگوریتم و افزایش میزان انتظار احتمال رخداد لگاریتمی در هر مرحله افزایش خواهد یافت یا تا در نهایت با رسیدن به یک بیشینه محلی ثابت خواهد شد.

^۱ Log-likelihood



شکل ۱۴ - مجموعه داده های آزمایشی و اجرای ۲۰ تکرار از الگوریتم تا همگرایی



شکل ۱۵ - رشد احتمال رخداد لگاریتمی به ازای هر تکرار از الگوریتم

۴-۳- مزایا و معایب

در این بخش تعدادی از مزایا و معایب الگوریتم بیشینه‌سازی انتظار را بررسی می‌کنیم.

مزایا:

- تضمین می‌کند که در هر دور پیمایش الگوریتم احتمال رخداد لگاریتمی افزایش پیدا کند.
- در عمل به خوبی کار می‌کند و می‌تواند الگوها و توزیع های آماری را به خوبی شناسایی کند.

معایب:

- تضمین نمی‌کند که لزوماً بیشینه احتمال رخداد را بدهد و ممکن است در بیشینه‌های محلی گیر کند.
- ممکن است همگرایی آن بسیار کند باشد و مراحل الگوریتم و هزینه زمانی زیاد شود.
- نیازمند هر دو احتمالات رو به جلو و رو به عقب می‌باشد.

فصل چهارم

کاربرد های EM

کاربرد های EM

امروزه به علت ویژگی های این الگوریتم همانطور که گفته شد - ویژگی هایی مانند خوشه بندی نرم و توانایی تشخیص الگو های پیچیده و ... - این الگوریتم در رشته ها و کاربرد های متنوع ای مورد استفاده قرار می گیرد.

۴-۱- کاربرد در رشته های مختلف

به طور مثال از این الگوریتم در کاربرد های زیر استفاده می شود:

۱. در علوم و مهندسی کامپیوتر برای کار هایی از قبیل فشرده سازی تصویر، خوشه بندی داده ها، پردازش زبان طبیعی و
۲. با توانایی مقابله با داده های از دست رفته و مشاهده متغیرهای ناشناخته ، EM در حال تبدیل شدن به ابزاری مفید برای قیمت گذاری و مدیریت ریسک نمونه کارها است.
۳. الگوریتم EM همچنین به طور گسترده ای در باز سازی تصویر پزشکی ، به ویژه در توموگرافی انتشار پوزیترون ، توموگرافی رایانه ای انتشار تک فوتون ، و توموگرافی رایانه ای اشعه ایکس استفاده می شود .

۴-۲- بررسی مثالی ساده :

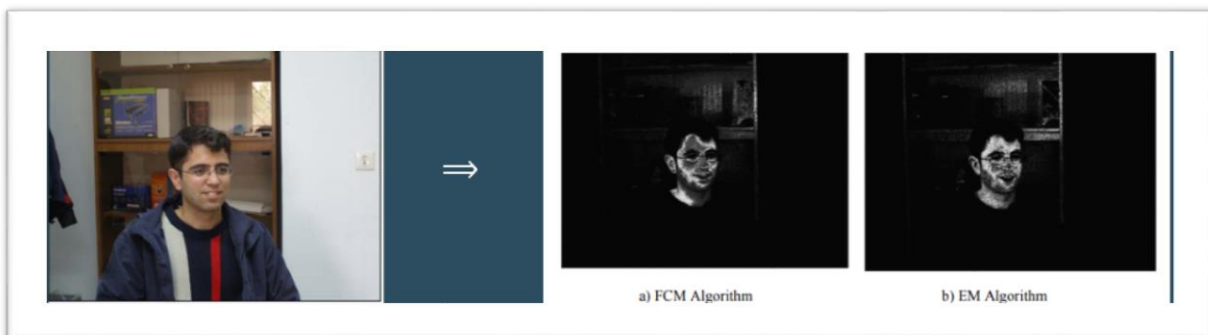
همانطور که در بخش قبل گفته شد. این الگوریتم برای کار هایی از قبیل فشرده سازی تصاویر استفاده می شود.

در تصاویر زیر یک مورد از این فشرده سازی ها قابل مشاهده است. همانطور که مشاهده می شود الگوریتم با تشخیص بخش های مخلف عکس قادر به کاهش سایز تصویر شده و خروجی تغییر چندانی با ورودی ندارد حال آنکه ساز آن به شدت کمتر می شود.



شکل ۱۶ - خروجی فشرده سازی

علاوه بر شکل ۱۷ در شکل ۱۸ نیز شاهد یک مقایسه بین الگوریتم EM و الگوریتم فازی FCM هستیم. در این تصویر گویا EM بهتر از FCM عمل کرده است. اما با توجه به این که هر دو این الگوریتم ها در پردازش تصاویر پزشکی همچنین در فشرده سازی تصاویر استفاده می شوند، مقایسه بین این دو الگوریتم خود نیازمند مقاله ای دیگر است.



شکل ۱۷ - مقایسه عملکرد EM و FCM

فصل پنجم

جمع‌بندی و نتیجه‌گیری و پیشنهادات

جمع‌بندی و نتیجه‌گیری

در این مقاله ابتدا به بیان اهمیت علم آمار و احتمالات در الگوریتم‌های خوشه‌بندی و چه بسا دانش هوش مصنوعی پرداختیم. در ادامه پس از معرفی الگوریتم EM به عنوان یک مثال از الگوریتم‌های مبتنی بر احتمالات به توضیح چگونگی عملکرد و پیاده‌سازی آن پرداختیم. در پایان به کاربرد‌های این الگوریتم رسیدیم و دانستیم که این الگوریتم در زمینه‌های مختلف راهگشای ما ست و میتوان در علوم مختلف از آن بهره جست.

پیشنهادهای

در انتها پیشنهاد می‌دهیم پس از تسلط بر این الگوریتم به مقایسه آن با الگوریتم FCM پردازید. چرا که این دو الگوریتم شباهت و تفاوت‌های جالبی با هم داشته و همچنین در عرصه‌ها مختلفی استفاده می‌شوند.

منابع و مراجع

- 1- Comparison between EM and FCM algorithms in skin tone extraction. Hu, Y., Wai, L. and Chee, S., 2012, November.
- 2- Clustering algorithms for scene classification—A performance comparison between K-means; fuzzyc-means and GMM. In Proc. IEEE Image Electron. Vis. Comput. Workshop (pp. 1-6).
- 3- CSC 411: Lecture 13: Mixtures of Gaussians and EM, Richard Zemel, Raquel Urtasun and Sanja Fidler Ravanbakhsh, E., Rezaei, M., Namjoo, E. and Choobdar.
- 4- https://en.wikipedia.org/wiki/Expectation%E2%80%93maximization_algorithm#:~:text=One%20of%20the%20earliest%20is,L%C3%B6f%20and%20Anders%20Martin%2DL%C3%B6f
- 5- <https://towardsdatascience.com/expectation-maximization-explained-c82f5ed438e5>
- 6- <https://medium.com/analytics-vidhya/expectation-maximization-algorithm-step-by-step-30157192de9f>
- 7- <https://github.com/volflow/Expectation-Maximization>