((بررسي روش هاي مربوط به شبكه هاي بيزين و كاربردهاي آن))

مرجان نيتي

Marjan_Niyati@yahoo.com

چكيده :

شبكه بيزين يك ابزار نسبتا جديد براي شناسايي (هويت) روابط احتمالي به منظور پيشگويي يا ارزيابي كلاس عضويت است . در اين مقاله به بررسي روش هاي به كار برده شده در شبكه بيزين و كاربردهاي آن خواهيم پرداخت .

۱- **مقدمه** :شبكه هاي بيزين در زمينه استدلال احتمالي به طور گسترده مورد استفاده قرار مي گيرند . **[۳۰][۲۲][۱۵]**شبكه هاي بيزين به درخت متصل بر روي احتمالات استدلال شده تبديل مي شوند . اخيرا شبكه هاي بيزين به((تجزيه زيرگراف اصلي ماكزيمم درخت متصل))[۱] تبديل مي شوند و بيشتر از درخت هاي متصل كاربرد دارند .**[۱۵]** اولين شبكه بيزين بر اساس دانش خبره ايجاد شد و عموما به صورت آشكار با مقادير اوليه قابل قبول و روابط ما بين متغيرها توزيع مي شود.**[۳]** شبكه هاي بيزين در مسائل دنياي واقعي بسيار كاربرد دارند . **[۲۲]** در چندين سال پيش شبكه هاي بيزين توسط افــــــراد

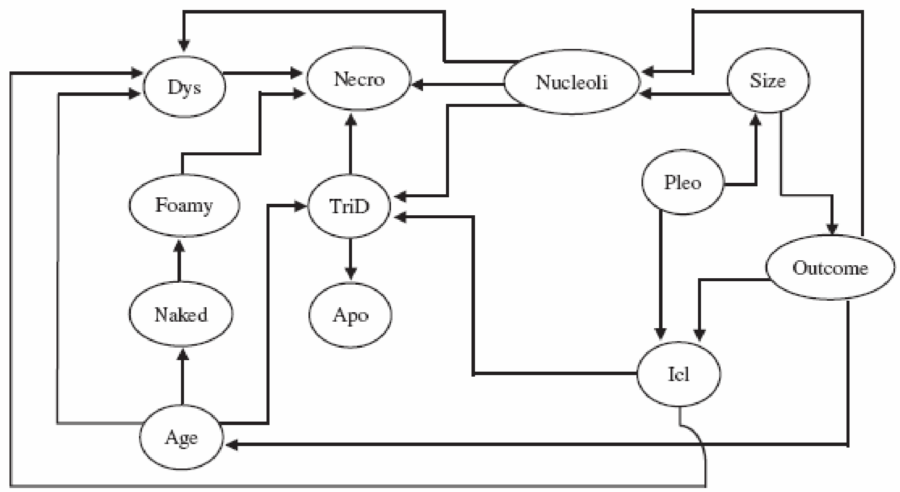
Herskovits, 1992; Heckerman et al., 1995; Heckerman, 1997; Neapolitan, 2003; Pearl, 1988)

مورد توجه قرار گرفتند و به عنوان گروههاي زيست شناسي در روش هاي شبكه هاي ژني توسط افرادي به كار گرفته شدند .**[۱۲]** شبكه بيزين يك مدل گرافيكي براي نمايش احتمالات ما بين متغيرهاي مورد نظر مي باشد .**[۷][۵]**. از طرفي شبكه هاي بيزين روشي براي نمايش توزيع احتمالي پيوسته بزرگ به صورت نمايي و روش فشرده مي باشند كه اجازه محاسبات احتمالي به طور موثر را مي دهند . آنها از ساختار مدل گرافيكي براي ضوابط مستقل ما بين متغير هاي تصادفي استفاده مي كنند . شبكه هاي بيزين اغلب براي شرايط مدل احتمالي استفاده مي شوند و به استدلالهاي تحت شرايط نامشخص (احتمالي ، عدم قطعيت) كمك مي كنند . **[۱۱]** اين شبكه شامل بخش كيفي (مدل ساختاري) است كه نمايش بصري از فعل و انفعالات در ميان متغيرها و بخش كمي (مجموعه اي از مشخصات احتمال محلي) را فراهم مي كند ، كه مجاز به استنتاج احتمالات و اندازه گيري عددي است كه متغيرها يا مجموعه اي از متغيرها را تحت تاثير قرار مي دهد .بخش كيفي به صورت توزيع احتمالي پيوسته كه منحصر به فرد مي باشد بر روي كليه متغيرها تعريف مي شود .**[۵]**به عبارت ديگر شبكه بيزين يك گراف جهت دار غير حلقوي است **[۵][۱۰][۲۰][۲۹] [۷]**كه شامل موارد زير مي باشد : **[۵]**

- نودها (دواير كوچك) : براي نمايش متغيرهاي تصادفي ، كمانها (پيكانهاي نوك تيز) براي نمايش روابط احتمالي ما بين متغيرها و
- براي هر نود توزيع احتمال محلي وجود دارد كه به نود وابسته است و از وضعيت والدين مستقل مي باشد . شكل زير يك مثال از شبكه بيزين و جدول احتمال شرطي متناظر (كه اين شبكه بيزين نتايج حاصل از تركيب شبكه بيزين با الگوريتم حريصانه با ۶۹۲ مجموعه از داده هاي حاصل از سرطان سينه مي باشد) را نمايش مي دهد .**[۵]**

---

[۱] Maximal Prime Subgraph Decomposition Junction Tree (MPD-JT)



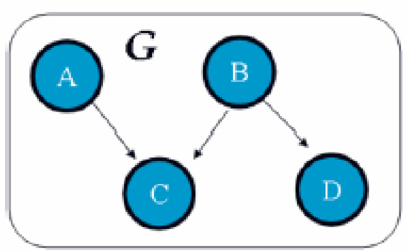
P(Outcome | Size)	
P(out = 0 | size = 0)	0.926
P(out = 1 | size = 0)	0.074
P(out = 0 | size = 1)	0.146
P(out = 1 | size = 1)	0.854

يك مثال از شبكه بيزين با جدول احتمال شرطي متناظر

از فوايد بزرگ اين مدل مي توان گفت كه اين مدل به ما اجازه توزيـع احتمـالي پيوسـته بـه صـورت فـشرده و مقرون به صرفه با استفاده از شرط مستقل در معادله زير را مي دهد :[۵]

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P(X_i|\mathrm{Pa}(X_i))$$

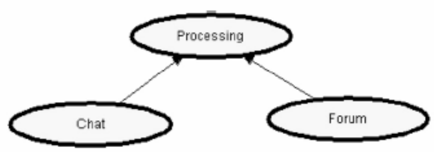
$\mathrm{Pa}(X_i)$ براي نمايش مجموعه نودهاي والد از X_i و كمانهايي با اشاره گر به كار مي رود . نمونه اي از گراف جهت دار غير حلقوي در شكل زير نمايش داده شده است :[۱۰]



گراف جهت دار غير حلقوي

Xو Y دو متغير شرطي مستقل با يك مجموعه متغير Z مي باشد P توزيع احتمال كه بـه صـورت $X \perp\!\!\!\perp_P Y|\mathbf{Z}$ ؛ مي باشد اگر شرط $P(X,Y|Z) = P(X|Z)P(Y|Z)$ بر قرار باشد . شرايط مـستقل بـه صـورت $X \not\perp\!\!\!\perp_P Y|\mathbf{Z} \equiv \neg X \perp\!\!\!\perp_P Y|\mathbf{Z}$ تعريف مي شود . اگر V مجموعه متناهي غير تهي از متغيرهاي تصادفي باشد در اين صورت شبكه بيزين براي Vيك جفت $<G,\theta>$ مي باشد كه G گراف جهت دار غير حلقوي است كه نودها متناظر با متغيرهاي تـصادفي در Vو θ پارامترهـاي ويـژه توزيـع احتمـال شـرطي بـراي هـر نـود X بـه عنـوان والـد در G بـه صـورت $p(X|Pa_G(X))$ مي باشد . [۱۲]براي غلبه بر مشكلات شرايط پيچيدگي محاسبات چندين تقريـب ماننـد مرتـب سازي هاي پيشين بر روي حوزه اي از ويژگيها بر روي تعداد محدودي از ساختارهاي بيزين مورد يادگيري قرار گرفته است يا استفاده از روش هايي كه منجر به كاهش فضاي حالت در مساله مي شوند . شـبكه هـاي بيـزين براي اهداف دسته بندي ساخته مي شوند و امكان اضافه كردن قيود با بازده محاسبات اضـافي كـه مـي توانـد

قابل اشتقاق باشد را بهبود مي بخشند . مرتب سازي متغيرها مي تواند توسط يك فرد خبره انجام پـذيرد امـا اين روش چندان مفيد نيست و با روش هاي يادگيري ماشين و داده كاوي انجام مي پـذيرد . يـادگيري شـبكه هاي بيزين الگوريتم هايي هستند كه اغلب براي مرتب سازي متغيرهاي تصادفي يا مرتـب سـازي متغيرهـا بـا جزييات به كار مي روند . الگوريتم هاي رده بندي ويژگي ها مي تواند براي ويژگي هاي مرتبط و قابل مقايسه تعريف شود . اين الگوريتم ها مي توانند براي يادگيري دسته بندي كننـده شـبكه بيـزين بـه كـار رونـد . [٦] الگوريتم بهينه سازي بيزين توسط Pelikan پيشنهاد داده شد . الگوريتم بهينه سازي بيزين [٢] مـدل احتمـالي است كه بر اساس شبكه هاي بيزين ساخته شده اند . شبكه هاي بيـزين در الگـوريتم بهينـه سـازي بيـزين در تركيب نودها و جهت لينك ها به كار مي روند .[١] به طور خلاصه مي توان گفت شبكه بيزين ، نمايش بامعني روابط نامشخص ما بين پارامتر ها در يك حوزه مي باشد . شبكه بيزين گراف جهت دار غيـر حلقـوي از نودهـا براي نمايش متغيرهاي تصادفي و كمانها براي نمايش روابـط احتمـالي مـا بـين متغيرهـا بـه شـمار مـي رود . (Jensen, 1996) . فقدان لبه ها در شبكه بيزين نشانگر جملات مستقل از هم مي باشد . رمز گـذاري شـبكه بيزين مطابق با جملات مستقل در هر متغير تصادفي است . يك متغير مـستقل بـه صـورت صـعودي وضـعيت والدين شبكه را نشان مي دهد . (Pearl, 1988) . همچنين شبكه بيزين براي نمايش توزيع احتمالي ويـژه و اتصال توزيع بر روي همه متغيرها به صورت نودها در گراف نمايش داده مي شود . اين توزيع با يك مجموعه از جدول احتمال شرطي [٣] مشخص مي شود . هر نود به جدول احتمال شرطي منتسب شـده و توسـط اطلاعـات احتمالي كمي مشخص مي گردد. همانند جـدولي احتمـالات در وضـعيت ممكـن از نـود در تركيـب ممكـن از والدينش مشخص مي گردد . براي نودهاي بدون والد احتمالات بر روي نودهاي ديگر بـدون قيـد و شـرط مـي باشند كه اين نودها احتمالات اوليه بر روي متغيرها ناميده مي شوند . در شكل زيـر يـك شـبكه بيـزين سـاده نمايش داده شده است . [٨]



Forum	Labelled	Forum
replies messag...	0.3	
reads messages	0.2	
posts messages	0.5	
no participation	0	

Chat	Labelled	Chat
participates	0.33333	
listens	0.33333	
no participation	0.33333	

Processing	Labelled	Processing										
Forum	replies messages			reads messages			posts messages			no participation		
Chat	partici...	listens	no par...	partici...	listens	no par...	partici...	listens	no par...	partici...	listens	no par...
active	0.85	0.7	0.55	0.75	0.6	0.45	1	0.85	0.7	0.65	0.5	0.35
reflecive	0.15	0.3	0.45	0.25	0.4	0.55	0	0.15	0.3	0.35	0.5	0.65

١- شبكه بيزين ساده

اين مدل شبكه اي روابط ما بين مشاركت دانش آموزان در چت (گفتگوي اينترنتي) و تبادل نظرو پردازش را نشان مي دهد . شبكه بيزين در اين مدل شامل سه نود پردازش ، تبادل نظر و چت مي باشـد .[٨] مـدل رياضـي شـبكه بيزين تئوري بيز است كه در معادله زير نشان داده شده است . [٨] [٢٦]

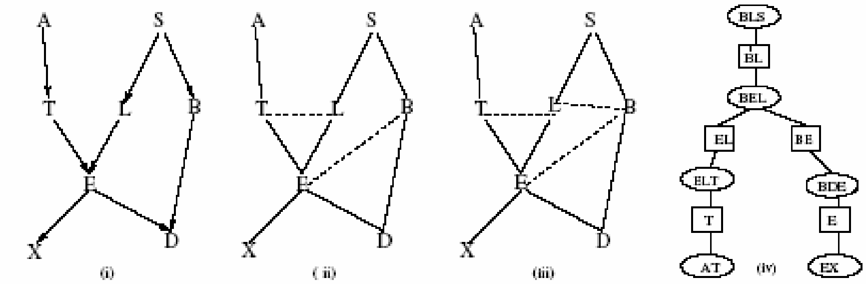
$$P(A/B) = \frac{P(B/A)P(A)}{P(B)}$$

[٢] BOA(Bayesian optimization algorithms)
[٣] CPT(conditional probability tables)

تئوري بيز روابط شروط و احتمالات حاشيه اي است . توزيع احتمالي شرطي يك متغير تـصادفي از A مـي باشـد . توزيع احتمال شرطي A به شرط B معادل با احتمال Bبه شرط A ضرب در احتمال A ، تقسيم بر احتمال B مي باشد . [۸] اگر بخواهيم تعريف ديگري از شبكه هاي بيزين داشته باشيم بـه ايـن ترتيـب اسـت كـه شـبكه بيـزين مجموعه اي از متغير گسسته محدود به صورت $V = \{A_i | i = 1, \ldots, n\}$ مي باشد كه بـه صـورت يـك گـراف جهت دار غير حلقوي با علامت $\mathscr{D}$ مشخص مي شود ، كه يك مجموعه از توزيع احتمال شرطي بـه آن اضـافه مـي شود . هر راس در $\mathscr{D}$ با متغير V تناظر يك به يك دارد كه V (رئوس ضوابط ، نـود و متغيرهـايي اسـت كـه بـا يكديگر تبادل پذير مي باشند) . والدين هر نود A_i در $\mathscr{D}$ با π_{A_i} مشخص مي شود . هر متغير $A_{si} \in V$ به تعيين ويژگــي هــايي از جــدول احتمــال شــرطي $p(A_i|\pi_{A_i})$ مــي پــردازد كــه توزيــع احتمــال پيوســته آن $p(V) = \prod_{A_i \in V} p(A_i|\pi_{A_i})$ مي باشد . به طور خلاصه $Y \Rightarrow \Rightarrow X|Z$ و X و Y و Z زيرمجموعـه جـدا از V مي باشند كه X و Z شرايط مستقل از Yمي باشند . اگر $XYZ = V$ باشـد در ايـن صـورت $X \cup Y \cup Z$ خواهد شد . گراف جهت دار غير حلقوي در شبكه هاي بيزين قابل تبديل به درخت متصل براي اسـتدلال احتمـالي است . اين تبديل شامل دو عملكرد مهم يكي درستي و نادرستي اعمال را مورد تجزيه و تحليل قرار دادن [٤] و ديگري مثلث بندي يا سه گوش بندي [5] مي باشد . تجزيه و تحليـل گـراف در گـراف جهـت دار غيـر حلقـوي $\mathscr{D}$ بـا $M^{\mathscr{D}}$ مشخص مي شود كه يك گراف غير جهت دار با اتصال هر جفت از نودها به فرزند كه در حـال حاضـر در گـراف $\mathscr{D}$ متصل نيست بدست مي آيد و سپس جهت كمان به سمت تمامي لبه ها جهت دار مي گردد . $M^{\mathscr{D}}$ با اضافه شـدن به وتر يا قوس [٦] كه لبه پر ناميده مي شود مثلث بندي مي گردد . مثلث بندي مينيمم است اگر حذف يك لبه پـر در نتيجه يك گراف غير مثلث بندي شده صورت گيرد . نتايج گراف مثلث بندي شده با $T^{\mathscr{D}}$ مـشخص مـي گـردد . درخت متصل $\mathscr{T}$ با مشخص كردن تمامي دسته ها مانند (زير گراف كامل ماكزيمم) در $T^{\mathscr{D}}$ و مرتب كردن آنهـا به عنوان نودهاي درخت ساخته مي شود . كه هر باند C مستلزم مسير ما بين دو باند C' and C'' در $\mathscr{T}$ كـه $C' \cap C''$ مي باشد . تمامي درخت هاي متصل چند گانه ممكن است از $T^{\mathscr{D}}$ ايجاد شوند ، كه وابـسته بـه مـنظم كردن باند ها براي فصل مشترك (تقاطع) اجرا مي باشد . اگر باندهاي C' and C'' با يـك لبـه در $\mathscr{T}$ متـصل شــوند ايــن لبــه بــه صــورت $C' \cap C''$ برچــسب گــذاري مــي شــود . مجموعــه $\mathscr{S}^{\mathscr{T}}$ بــه صــورت $\mathscr{S}^{\mathscr{T}} = \{C \cap C' \mid \text{there is an edge between } C \text{ and } C'\}$ تعريف شده و مجموعه منفصل از $\mathscr{T}$ ناميده مي شوند . كه هر عنصر $\mathscr{S}$ يك جدا كننده در $\mathscr{T}$ ناميده مي شود . به شـبكه بيـزين "Asia" در شـكل زير توجه كنيد گراف تجزيه و تحليل گر $M^{\mathscr{D}}$ و گـراف مثلـث بنـدي شـده $T^{\mathscr{D}}$ در شـكل (ii) and (iii) بـه ترتيب نشان داده شده اند لبه هاي $(T, L)^1$ و (E, B) در $M^{\mathscr{D}}$ به صورت نقطه چين در حين تجزيه و تحليل به شكل (ii) اضافه مي شوند . لبه متصل (L, B) به صورت نقطه چين در (iii) نشان داده شده است كـه در طـي

[٤] moralization
[٥] triangulation
[٦] chord

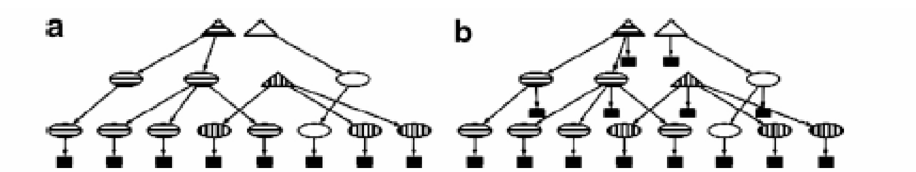
مثلـــــــــــــــث بنـــــــــــــدي اضـــــــــــــافه مـــــــــــــي شـــــــــــود .



[۱۵

(i) - گراف جهت دار غير حلقوي در شبكه بيزين "Asia" : (ii) - گراف تجزيه و تحليل گر $M^{\mathcal{D}}$ در گراف جهت دار غير حلقوي $\mathcal{D}$. (iii) - گراف مثلث بندي $T^{\mathcal{D}}$ در گراف جهت دار غير حلقوي $\mathcal{D}$. (iv) - درخت متصل ايجاد شده $\mathcal{T}$ از $T^{\mathcal{D}}$.

شش باند در گراف مثلث بندي $T^{\mathcal{D}}$ وجود دارد و آنها به صورت يك درخت متصل $\mathcal{T}$ در شكل (iv) مـنظم شـده اند . **[۱۵]** شكل زير ساختاري از شبكه هاي بيزين با ساختار درخت پويا را نمايش مي دهد .

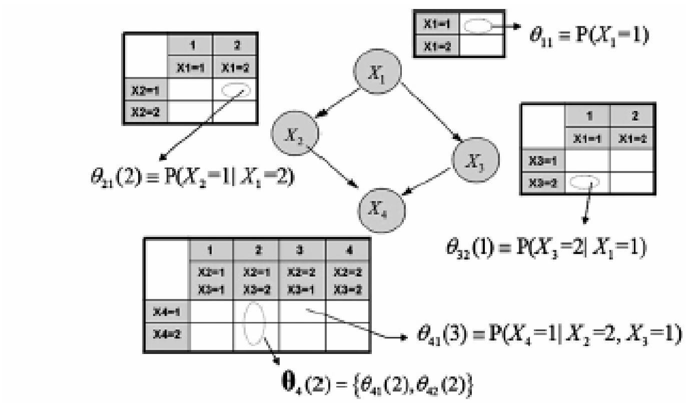


a – شبكه بيزين با ساختار درختي پويا **b** – درخت پويا . شبكه بيزين با ساختار درختي پويا شامل زير درخت هايي از جنگل ها و هر قطعه يك سيگنال تك بعدي به شكل نواحي علامتدار شده با سايه مشخص دايره اي و نودهاي مربعي براي متغيرهاي مخفي و مثلث ها براي نمايش ريشه ها مي باشند .

شبكه بيزين با ساختار درختي پويا مي تواند به صورت ساختار درختي براي توسعه و پردازش در پردازش تـصوير و بينايي ماشين به كار گرفته شود . براي مثال ساختار درختي از شبكه به صورت چند مقياسي قطعه بندي شده و در شبيه سازي تصوير و كاربردهاي پزشكي به كار گرفته مي شود . **[۲٤]** تعاريف متعـددي بـراي شـبكه هـاي بيـزين وجود دارد در تعريف ديگر ، شبكه بيزين B از يك جفت $B=(G,D)$ تشكيل شده است كـه G گـراف جهـت دار غير حلقوي و D مجموعه اي از توزيع احتمالي شرطي متناظر با G مي باشـد . نودهـاي گـراف در شـبكه هـاي بيزين متناظر با متغيرهاي تصادفي ، سيگنالها ، اندازه گيري (سنجش) [۷] يا مشاهده (شهودي) [۸] مي باشد . ضوابط سيگنال يا متغيرهاي تصادفي به نود شبكه بيزين مربوط مي شوند و به نامهاي سنجش يا شهود دربافتار آزمايـشات به صورت تكراري مورد استفاده قرار مي گيرند . **[۱٦]** گراف جهت دار غير حلقوي $\mathcal{G}$، بـا يـك مجموعـه متغيـر بـه صورت $X_1, X_2, \ldots, X_n$ و روابط ما بين آنها نمايش داده مي شود . هر نود در گراف جهت دار غيـر حلقـوي $\mathcal{G}$، بـه عنوان متغير و جهت كمانهاي بين نودها روابط ما بين وابسته به والدين را نمايش مي دهد . در شكل زير در هر نـود يك توزيع احتمال شرطي براي متغير و مقادير والدين وجود دارد . **[۲۱]**

---

۷ measurements
۸ observations



توزيع احتمال شرطي در شبكه بيزين

توزيع احتمال پيوسته در شبكه بيزين به صورت زير مي باشد :[۲۷][۲۱]

$$p(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n|\mathscr{G}) = p(x_1, x_2, \ldots, x_n|\mathscr{G}) = \prod_{i=1}^{n} p(x_i \mid x_{\mathrm{pa}(i)}, \mathscr{G}).$$

$p(x_1, x_2, \ldots, x_n|\mathscr{G})$ نمايش احتمـالي تركيـب ويـژه اي از $x_1, x_2, \ldots, x_n$ در متغيرهـاي $X_1, X_2, \ldots, X_n$ مـي باشد و $x_{\mathrm{pa}(i)}$ به عنوان يك بردار جهت والدين X_i، است كه در گراف جهت دار غير حلقوي $\mathscr{G}$ نمايش داده مي شود . شبكه بيزين ساختار محلي دارد اگر هر نود تاثير متقابل به نود والد داشته باشد . براي متغيرهاي تصادفي گسـسته توزيـع احتمـال شـرطي يـك مجموعـه از جـداول مـي باشـد كـه بـه توزيـع چندگانـه بـه پارامترهـاي $\theta = \{\theta_1, \theta_2, \ldots, \theta_n\} = \{[\theta_{ik}(j)]_{k=1}^{r_i}\}_{j=1}^{q_i}$ مـي پـردازد بـه طـوري كـه $i = 1 \ldots n$ مـي باشـد و هـر متغيـر $X_i \in \mathbf{X};\ k = 1 \ldots r_i$ داراي مقدار $X_i; j = 1 \ldots q_i$ با وضعيت مقادير اعتبار براي متغيرهاي والدين $X_i\ (x_{\mathrm{pa}(i)})$. مي باشد . باي توزيع گسسته توزيع احتمال پيوسته در گراف جهت دار غير حلقوي به صورت زير مي باشد :[۲۱]

$$p(\mathbf{x}|\boldsymbol{\theta}, \mathscr{G}) = \prod_{i=1}^{n} p(x_i|x_{\mathrm{pa}(i)}, \theta_i, \mathscr{G}) = \prod_{i=1}^{n} \theta_{ik}(j)$$

۲- دسته بندي الگوريتم هاي بيزين با روش هاي مختلف

۲-۱- الگوريتم هاي بهينه سازي بيزين :

محاسبات تكاملي الگوريتم هاي توزيع تخمين EDAs[۹] ، اولين بار توسط Muhlenbein معرفي شدند . EDAs الگوريتم هاي ژنتيك ساختاري مدل احتمالي به عبارتي PMBGAs[۱۰] ناميده مي شوند و براي تـشخيص بافتـاري جستجوي تكاملي بزرگ بدون دانش اوليه و ارتباط با مساله مورد مطالعه قرار مي گيرنـد . الگـوريتم هـاي EDAs اگر چه مشابه الگوريتم هاي ژنتيك مي باشند اما به مراتب همانند آنها عمل نمي كنند . عملگرهـاي ژنتيـك صـريح مانند تقاطع [۱۱] و جهش مي باشند . در عوض آنها مدلهاي احتمالي از مجموعه اي از راه حل هاي اميدوار كننـده – مقادير شايستگي از رشته اي طويل (براي ماكزيمم كردن مساله) يا كوتاهتر (براي مينيمم كردن مـساله) سـاخته مي شوند ، و اين مدل بوجود آمده براي توليد رشته بعدي به كار برده مي شود .عموما EDAs به صورت دنبالـه اي تكراري از راه حل هاي بهينه كه در زير به آن اشاره شده بدست مي آيد . [۱]

۲-۱-۱- يك جمعيت از N رشته به طور تصادفي مقدار دهي اوليه مي شود .

[۹] Estimation of distribution algorithms
[۱۰] Probabilistic model-building genetic algorithms
[۱۱] crossovers

۲-۱-۲- ر شته هاي مرتب شده توسط شايستگي و انتخاب $M\ (M < N)$ بـا شايـستگي بـالا (بـراي مـساله ماكزيمم سازي) انتخاب مي شود .

۲-۱-۳- يك مدل احتمالي بر اساس رشته هاي انتخاب شده ايجاد مي شود .

۲-۱-۴- يك نسل جديد از N رشته بر اساس مدل ايجاد مي شود .

۲-۱-۵- اگر شرط خاتمه صورت نگرفت به مرحله ۲-۱-۲ رفته و تكرار الگوريتم .

هدف اصلي EDAs اين است اگر توانستيم چند باياس در توزيع شايستگي از رشته جمعيت تـشخيص دهـيم در اين صورت راه حل هاي خوب در توليد نسل از يك مدل بر اساس انتخـاب از يـك زيـر جمعيـت را حـدس خواهيم زد . روش هاي مختلف پيشنهاد شده براي EDAs مانند يادگيري افزايـشي بـر اسـاس جمعيـت [۱۲] ، الگوريتم توزيع حاشيه اي جمعي (عمـومي) [۱۳] ، الگـوريتم توزيـع حاشـيه اي دو تـايي [۱٤] و غيـره مـي باشـد . الگوريتم ژنتيك فشرده و الگوريتم ژنتيك فشرده پيشرفته [15] به عنوان يادگيري افزايـشي بـر اسـاس جمعيـت پيشنهاد داده شدند . هر نود به صورت $X_i\ (i = 0, \ldots, l-1)$ مي باشد كه l طول رشته و شبكه با متغير هـاي تصادفي نمايش داده مي شود . احتمال مكان هندسي si مقدار l مي باشد . لينك جهت دار از نود i به نود j به منظور نمايش نود i به نود والد j ايجاد مي شود . توزيع احتمالي هر متغير تصادفي X_i مـي توانـد بـه صـورت $p(X_i|\prod_{X_i})$ نوشته شود . $\prod_{X_i}$ مجموعه والد از X_i: مي باشد . مجموعه اي از نودهايي هـستند كـه بـه نـود i متصل شده اند . به طور كلي توزيع احتمالي هر رشته X با اتصال به توزيع زير تعريف مي شود :[۱]

$$p(X) = \prod_{i=0}^{n-1} p\left(X_i \middle| \prod_{X_i}\right)$$

مدل ساختاري و فرآيند جستجوي الگوريتم بهينه سازي بيزين مطابق با روال زير مي باشد :

۲-۱-۶- مجموعه $t \leftarrow 0$ به طور تصادفي براي توليد نسل از جمعيت اوليه $P(0)$ انتخاب مي شود .

۲-۱-۷- يك مجموعه رشته اميدبخش $S(t)$ از $P(t)$ انتخاب مي شود .

۲-۱-۸- شبكه B با استفاده از پارامتر انتخاب و قيود ساخته مي شود .

۲-۱-۹- يك مجموعه از رشته جديد $O(t)$ توليد شده مطابق با توزيع احتمال اتصالي توسـط B رمزگـذاري مي شود .

۲-۱-۱۰- ايجاد يك جمعيت اوليه $P(t+1)$ با تعداد رشته هاي $P(t)$ به همراه رشـته هـايي از $O(t)$، و مجموعه $t \leftarrow t+1$ جايگزين مي شود .

۲-۱-۱۱- اگر شرط خاتمه نبود به مرحله ۲-۱-۷ برگشته و ادامه الگوريتم از سر گرفته مي شود .

مشابه الگوريتم هاي ژنتيكي جمعيت اوليه الگوريتم هاي بهينه سازي بيزين جمعيت اوليه به طور تصادفي از رشـته باينري توليد شده و يك زير مجموعه از راه حل هاي اميد بخش به نسبت بالا از مقادير شايستگي انتخاب مي شوند . كار شبكه بيزين B با استفاده از اين رشته ها و بر اساس يك پارامتر ويژه و يك مجموعه قيـود سـاخته مـي شـود . ارزيابي كيفيت شبكه BOA با پارامتر هدايت پذير بيزين به صورت زير تعريف مي شود . [۱]

$$p(D,B|\xi) = p(B|\xi)\prod_{i=0}^{n-1}\prod_{\pi_{X_i}}\frac{\Gamma(m'(\pi_{X_i}))}{\Gamma((m'(\pi_{X_i})+m(\pi_{X_i})))}\times\prod_{x_i}\frac{\Gamma((m'(x_i,\pi_{X_i})+m(x_i,\pi_{X_i})))}{\Gamma(m'(x_i,\pi_{X_i}))}$$

---

۱۲ PBIL (Population- Based incremental Learning)
۱۳ UMDA(univariate marginal distribution algorithms)
۱٤ BMDA(Bivariate marginal distribution algorithms)
۱٥ ECGA (extended compact GA)

matlab1.ir

D مجموعه اي از داده از راه حل هاي كارا (اميد بخش) مي باشد . $p(B|\xi)$ احتمال اوليه از شبكه B مـي باشـد . Γ تابع تعريف شده به صورت $\Gamma(a) = (a-1)!$ و $m(\pi_{X_i})$ نمايش تعداد داده ها ، $\prod_{X_i}$ برابـر بـا π_{X_i} در D مـي باشـــــــد . $m(x_i, \pi_{X_i})$ تعـــــــداد داده X_i از x_i و $\prod_{X_i}$ از π_{X_i} در D مــــــــي باشـــــــد . $(m(\pi_{X_i}) = \sum_{x_i} m(x_i, \pi_{X_i}))$. $m'(x_i, \pi_{X_i})$ با $p(B|\xi)$ از اطلاعات اوليه شبكه $m(x_i, \pi_{X_i})$ با توجه به مساله بدست مي آيد . براي احتمال اوليه $p(B|\xi)$ ، Heckerman تابع احتمال ساده مانند $p(B|\xi) = c\kappa^{\delta}$ را پيشنهاد كـرد . c ثابت نرماليزه شده ، $\kappa \in (0,1)$ پارامتر ثابت ديگر و δ ماكزيمم درجه شبكه (مانند ماكزيمم تعداد اتصالات بـه نود) با مقادير ثابت k مي باشد . براي انتخاب رشته ، الگوريتم بهينه سازي بيزين به صورت جستجوي حريـصانه بـا مينيمم كردن پارامتر تحت اين قيد اجرا مي شود . الگوريتم حريصانه به صورت زير اجرا مي شود .

۲-۱-۱۲- مقداردهي اوليه شبكه B .

۲-۱-۱۳- انتخاب همه اپراتورهاي گراف ساده كه مي تواند بر روي شبكه بدون نقض قيود اجرا شود .

۲-۱-۱۴- برداشتن عمليات به منظور افزايش امتياز شبكه .

۲-۱-۱۵- انجام عمليات در مرحله قبلي .

۲-۱-۱۶- اگر شبكه نتوانست طولاني باشد بدون نقض قيود به طور پيچيده گسترش يابد يا زماني كـه بـه تعـداد ماكزيمم فعل و انفعال رسيد خاتمه يابد .

۲-۱-۱۷- به مرحله ۲-۱-۱۳ برو .

نسل جديد از نمو نه هاي نسل بعدي انتخاب مي شود . و الگوريتم بهينه سازي بيزين به صورت زير انجام مي شود :

۲-۱-۱۸- تمام متغير هاي غير پردازشي علامت زده مي شوند .

۲-۱-۱۹- متغير غير پردازشي X_i با همه والدين در حال حاضر در پردازش برداشته مي شود .

۲-۱-۲۰- در مجموعه X_i ، x_i با احتمال $p(X_i = x_i|\prod_{X_i} = \pi_{X_i})$ مي باشد .

۲-۱-۲۱- X_i به عنوان پردازش علامت زده مي شود . [۱]

۲-۱-۲۲- اگر متغيرهاي تصادفي در چپ وجود داشت به مرحله ۲-۱-۱۹ برو.[۱]

۲-۲- الگوريتم بيزين ساده مثبت : متغيرها با حروف بزرگ و مقادير با حروف كوچك و بردارهـا بـا فونـت درشت نمايش داده مي شوند . در هر مساله دسته بندي بردار پيشگويي $\boldsymbol{X} = \{X_1, \ldots, X_n\}$ براي متغيرهاي تصادفي گسسته و C متغير كلاس تصادفي گسسته مي باشد . هر متغير تـصادفي X_i مـي تواند مقادير از ۱ تا r_i باشد و كلاس متغير C مي تواند دو مقدار ۱ براي نمونه هاي مثبت و ۰ بـراي نمونه هاي منفي باشد . مجموعه داده از دسته بندي كننده بـه صـورت $\mathscr{E}$ نمـايش داده مـي شـود و مشتمل بر يك مجموعه نمونه مثبت $(\mathscr{E}_1)$ و نمونه برچسب گذاري نشده $(\mathscr{E}_u)$, $\mathscr{E} = \mathscr{E}_1 \cup \mathscr{E}_u$ مي باشد . بيزين ساده مثبت در سال ۲۰۰۲ توسط Denis معرفي شد و سازگار با الگوريتم اسـتقرايي ساده بيز است كه در سال ۱۹۶۱ توسط Minsky براي بافتار يادگيري غير برچسب گـذاري مثبـت صورت گرفت . اين الگوريتم پارامترهاي مدل بيز ساده از مثالهاي مثبت و غير برچسب گذاري شـده را تخمين مي زند . قوانين بيز ارائه شده و مطابق با فرضيات مستقل شـروط مـا بـين تمـامي متغيرهـاي كلاس پيشگويي شده مي باشد كه براي نمونه x به صورت زير تعريف مي شود :[۴]

$$P(C = c|\boldsymbol{X} = \boldsymbol{x}) \propto P(C = c) \prod_{i=1}^{n} P(X_i = x_i|c)$$

پارامترهـاي مـورد نيـاز بـه صـورت دو كـلاس از مـدل بيـزين سـاده $P(C=1),\ P(X_i=j|C=1)$ و $P(X_i=j|\quad C=0)$ براي كليـه $i=1,\ldots,n$ و $j=1,\ldots,r_i-1$ بـه منظـور سـاده سـازي از احتمالات قبلي به صورت علايم $p,\ P(x_{ij}|1)$ و $P(x_{ij}|0))$ تعريف مي شود . در الگـوريتم بيـزين سـاده ايـن پارامترها از داده با ارزيابي احتمال ماكزيمم تخمين زده مي شوند ، اما بافتار يادگيري غير بـر چـسب گـذاري بـه دليل فقدان نمونه هاي منفي است كه غير عملي است و براي تخمين $P(x_{ij}|0)$ از P داده به كار مي رود . اگر داشته باشيم :

$$P(x_{ij}|0)=\frac{P(x_{ij})-P(x_{ij}|1)p}{1-p}$$

براي $P(x_{ij})$ براي $P(X_i=j)$ مد نظر است و مي توانيم $P(x_{ij}|0)$ براي P به صورت زير به كار بريم :

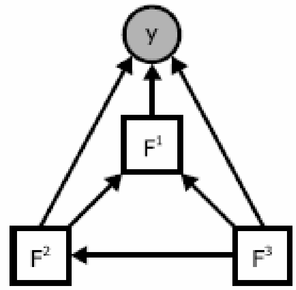
$$\frac{N_{ij\mathscr{E}_u}-P(x_{ij}|1)pN_{\mathscr{E}_u}}{(1-p)N_{\mathscr{E}_u}}$$

$N_{ij\mathscr{E}_u}$ مقدار نمونه غير برچسب گذاري $X_i=j$ و $N_{\mathscr{E}_u}$ و كاردينـاليتي مجموعـه مثالهـاي غيـر برچـسب گذاري شده است . مساله با اين ارزيابي كننده مي تواند منفي شود . حل مساله با جايگزيني تخمين منفي با صفر و نرماليزه كردن همه احتمالات براي هر متغير $X_i, \sum_{j=1}^{r_i}P(x_{ij}|0)=1$ پـيش بينـي مـي شـود . نرمـاليزه كردن و همبستگي لاپلاس ، همبستگي هموار مي باشد كه احتمالات ارزيابي شـده مـا بـين احتمـالات اوليـه در كلاس به صورت زير شيفت داده مي شود :**[٤]**

$$P(x_{ij}|0)=\frac{1+\max(0;R_i(j))\frac{1}{Z_i}}{r_i+(1-p)N_{\mathscr{E}_u}}$$

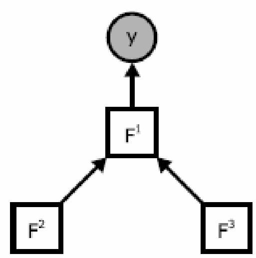
كـــــــه $R_i(j)=N_{ij\mathscr{E}_u}-P(x_{ij}|1)pN_{\mathscr{E}_u}$ مــــــي باشـــــــد و فـــــــاكتور نرمـــــــاليزه $\frac{N_{ij\mathscr{E}_u}-P(x_{ij}|1)pN_{\mathscr{E}_u}}{(1-p)N_{\mathscr{E}_u}}$ $Z_i=\sum_{j=1}^{r_i}\max(0;P(x_{ij}|0))$ اسـت كـه بـا اسـتفاده از معادلـه ارزيابي مي گردد . p نمي تواند از داده تخمين زده شود بنابراين كاربر بايد پارامتري را تعريف نمايد . **[٤]**

٢-٣- شبكه هاي بيزين پويا : شبكه هاي بيزين پويا مدل گرافيكي ويژ ه اي از يك مجموعـه از متغيـر هـاي تصادفي و وابسته براي نمايش نودها و لبه ها در گراف جهت دار غير حلقوي ميبا شند . در مدت زماني كه شـبكه هاي بيزين شرح داده مي شوند نودهايي را به عنوان علايم مربعي / مـدور بـراي نمـايش متغيرهـاي گسـسته يـا پيوسته يا تمايزات سايه دار يا غير سايه دار براي متغير هاي قابل مشاهده يا مخفي به كار مي بريم . بـراي مثـال شكل زير Y متغير تصادفي و F^1, F^2, F^3 مقادير مخفي و گسسته مي باشند . مجموعه لبه هـا و جهـت هـا در شكل زير نمايش داده شده است . **[١٣]**



مثالي از شبكه بيزين با احتمالات پيوسته بر روي Y, F^1, F^2, F^3 كه مي تواند فاكتورهايي بر روي
$p(Y, F^1, F^2, F^3) = p(Y|F^1, F^2, F^3)P(F^1|F^2, F^3)P(F^2|F^3)P(F^3).$

به واسطه آن احتمالات پيوسته به صورت $p(Y,F^1,F^2,F^3) = p(Y|F^1,F^2,F^3)P(F^1|F^2,F^3)P(F^2|F^3)P(F^3)$ مي توانند باشند . كليد اصلي در فاكتورگيري مي باشد و والدين به صورت $Pa(\,).$ نمايش داده ميشوند . براي نود Y ، $Pa(Y)$ را تعريف مي كنيم به صورت يك سري از مجموعه نودها بر روي Y مشروط شده است به همين علت مطابق شكل $Pa(Y) = \{F^1, F^2, F^3\}.$ را خواهيم داشت . در شكل زير تعدادي از فرضيات مستقل ايجاد شده است براي مثال Y جهت مستقل بر روي F^2 and F^3 مي باشد و احتمال آن $Pa(Y) = \{F^1\}$ است . توزيع پيوسته Y, F^1, F^2, F^3 مي تواند به صورت $p(Y=y,F^1,F^2,F^3) = p(Y=y|F^1)P(F^1|F^2,F^3)P(F^2)P(F^3)$ فاكتورگيري شود . **[۱۳]**



نمونه اي از شبكه بيزين كه مي تواند $p(Y=y,\ F^1,\ F^2,\ F^3) = p(Y=y|F^1)P(F^1|F^2, F^3)P(F^2)P(F^3)$ فاكتور گيري شود .

شبكه هاي بيزين پويا به صورت نمونه اي از تكرار شبكه هاي بيزين در هر لحظه با اضافه شدن كمانها به متغيرهاي پيوسته (متصل) در زمانهاي مختلف مي باشند . در واقع شبكه هاي بيزين پويا براي نمايش انواع وابسته يا توزيع متغيرهاي تصادفي به كار مي روند . **[۱۳]**

۲-۴- شبكه هاي بيزين گوسين : G گراف جهت دار غيرحلقوي با مجموعه رئوس $V(G)$ و مجموعه يالهاي $E(G)$ مي باشد . فرض مي كنيم $V(G) = [n] := \{1, 2, \ldots, n\}$ باشد . فرض مي كنيم رئوس به ترتيب عددي مرتب شده باشند . در اين صورت $i \to j \in E(G)$ مي باشد اگر و فقط اگر $i < j$. باشد . شبكه بيزين منتسب شده با اين گراف مي تواند با فرمول فاكتورگيري برگشت پذير يا با جملات مستقل شرطي مشخص گردد . ابتدا به نمايش فاكتورگيري برگشتي متمركز مي شويم و با استفاده از اشتقاق جبري به توصيف پارامترها مي پردازيم . سپس با معرفي قيود مستقل شرطي به از بين بردن قيود توليد شده جبري بر روي اين مدل مي پردازيم . $X = (X_1, \ldots, X_n)$ يك بردار تصادفي و $f(x)$ تابع

چگالي احتمال برروي اين بردار تصادفي مي باشد . تئوري بيز چگالي پيوسته اي است كه مي تواند با فرمول زيـر فاكتورگيري شود :

$$f(x)=\prod_{i=1}^{n} f_i(x_i|x_1,\ldots,x_{i-1}),$$

$f_i(x_i|x_1,\ldots,x_{i-1})$ علايم چگالي شرطي است كه $X_1=x_1,\ldots,X_{i-1}=x_{i-1}$ بـه شرط X_i مي باشـد . [۱۷] مشخـصات فـاكتورگيري برگـشتي مـدل گرافيكـي اسـت كـه هـر چگـالي شـرطي $f_i(x_i|x_1,\ldots,x_{i-1})$ فقـــــــــط وابـــــــــسته بـــــــــه والـــــــــد $\text{pa}(i)=\{j\in[n]\mid j\rightarrow i\in E(G)\}$ مي باشد . كه مي توانيم به صورت زير نمايش دهيم .

$$f_i(x_i|x_1,\ldots,x_{i-1})=f_i(x_i|x_{\text{pa}(i)})$$

تابع چگالي f متعلق به شبكه بيزين است اگر فاكتورگيري به صورت زير باشد :

$$f(x)=\prod_{i=1}^{n} f_i(x_i|x_{\text{pa}(i)})$$

دنباله اين پارامترها در شرايط گوسي بدست مي آيد كه ابتدا ما مستلزم چند متغير تصادفي گوسي براي صدا زدن مي باشيم . هر متغير تصادفي n بعـدي از X بـا بـردار ميـانگين μ و Σ مـاتريس كواريـانس محـدود مثبـت مشخص مي شود . در اين صورت تابع چگالي پيوسته X به صورت زير مي باشد :

$$f(x)=\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$$

$|\Sigma|$ دترمينان Σ مي باشد . گوسي چند متغيره خانواده اي از "Bell curve" از گوسين جامع (عمـومي) مي باشد و توزيع مهم در تئوري احتمال و چند متغيره در آمار است چرا كه تئوري محدود مركزي مي باشـد . X متغير تصادف n بعدي و $A\subseteq[n]$ مي باشد . در اين صورت خواهيم داشت $X_A=(X_a)_{a\in A}$. اگـر x يك بردار باشد x_A زير برداري از كه توسط A شاخص گذاري مي گردد . براي ماتريس Σ ، $\Sigma_{A,B}$ زير ماتريس هايي از Σ با مجموعه شـاخص رديفـي A و مجموعـه شـاخص سـتوني B مـي باشـد . در ايـن ميـان مشخصات خوب از متغير تصادفي گوسي حاشيه نگاري شده و شرايط مشخصات گوسي حفظ مي شود .

فرض كنيد $X\sim\mathcal{N}(\mu,\Sigma)$ و $A,B\subseteq[n]$ منفصل باشند آنگاه خواهيم داشت :

(1) $X_A\sim\mathcal{N}(\mu_A,\Sigma_{A,A})$, *and*
(2) $X_A|X_B=x_B\sim\mathcal{N}(\mu_A+\Sigma_{A,B}\Sigma_{B,B}^{-1}(x_B-\mu_B),\Sigma_{A,A}-\Sigma_{A,B}\Sigma_{B,B}^{-1}\Sigma_{B,A})$.

شبكه بيزين گوسي با انتساب به گراف جهت دار غير حلقوي ساخته مي شود كه مجاز به توزيع شـرطي گوسـين بـــــراي توزيـــــع $f(x_i|x_{\text{pa}(i)})$ مــــي گـــــرديم . ايـــــن توزيـــــع شـــــرطي بـــــه صـــــورت

$$X_j = \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} X_i + W_j$$

قابل بايافت مـي باشـد . $W_j \sim \mathcal{N}(\nu_j, \psi_j^2)$ و

X_i با شرط $i < j$ و λ_{ij} پارامتر رگراسيون مستقل مي باشد . تبديل خطي متغيرهاي تـصادفي گوسـي گوسين بوده و X متغير تصادفي گوسي مي باشد . X با ميانگين μ و ماتريس كواريانس Σ مشخص مي شود كه بايد به صورت توزيع شرطي محاسبه شوند . توصيفات برگشت پذير براي توزيع X_j ، متغيرهاي بـا اولويـت مقدم (پيشي) كه براي ميانگين و ماتريس كواريانس به صورت برگشت پذير مي باشد . [۱۷]

$$\mu_j = \mathbb{E}(X_j) = \mathbb{E}\Bigg(\sum_{i \in \mathrm{pa}(j)} \lambda_{ij} X_i + W_j \Bigg) = \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} \mu_i + \nu_j$$

اگر $k < j$ باشد در اين صورت كواريانس :

$$\begin{aligned}
\sigma_{kj} &= \mathbb{E}\big((X_k - \mu_k)(X_j - \mu_j)\big) \\
&= \mathbb{E}\Bigg((X_k - \mu_k) \Bigg(\sum_{i \in \mathrm{pa}(j)} \lambda_{ij}(X_i - \mu_i) + W_j - \nu_j \Bigg) \Bigg) \\
&= \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} \mathbb{E}\big((X_k - \mu_k)(X_i - \mu_i)\big) + \mathbb{E}\big((X_k - \mu_k)(W_j - \nu_j)\big) \\
&= \sum_{i \in \mathrm{pa}(j)} \lambda_{ij} \sigma_{ik}
\end{aligned}$$

و واريانس

$$\begin{aligned}
\sigma_{jj} &= \mathbb{E}\big((X_j - \mu_j)^2\big) \\
&= \mathbb{E}\Bigg(\Bigg(\sum_{i \in \mathrm{pa}(j)} \lambda_{ij}(X_i - \mu_i) + W_j - \nu_j \Bigg)^2 \Bigg) \\
&= \sum_{i \in \mathrm{pa}(j)} \sum_{k \in \mathrm{pa}(j)} \lambda_{ij} \lambda_{kj} \sigma_{ik} + \psi_j^2.
\end{aligned}$$

به اين صورت تعريف مي شود . اگر هيچ قيد براي بردار ν وجود نداشته باشد در اين صورت براي μ هيچ قيدي نخواهد بود . در اين صورت بايد به قيود بر روي ماتريس كواريانس Σ توجه كنـيم . اگـر فـرض كنـيم ψ_j^2 بدون قيد ايجاد شده است در اين صورت مي توانيم اصطلاحات نامنظم (پريشان) را بـا σ_{jj} بـا پـارامتر سـاده جديد a_j جايگزين نماييم . كه منجر به نمايش جبري براي مدل ما مي گردد كه *trek rule* ناميده مي شود . براي هر لبه $i \to j \in E(G)$ ، λ_{ij} و براي هر راس $a_i : i \in V(G)$ تعيين شده است . فر

مي كنيم رئوس به ترتيب عددي مرتب شده اند كه $i \to j \in E(G)$ اگر و فقط اگر $i < j$ باشد . جفت لبه ها با عنوان يكسان $i \to k,\ j \to k$ تصادم مي كنند و برا هر جفت راس i, j يك $T(i, j)$ مجموعه اي از مسير ساده P در G از i به j وجود دارد كه در p تصادم نمي كند . مسيرهاي تصادم trek ناميده مي شوند . [۱۷]

۲-۵- آموزش بيزين با شبكه عصبي (شبكه هاي عصب بيزين) : فرض مي كنيم پارامترهاي احتمالي مشخص با داده و مدل ساختاربندي شده باشند . پارامترهاي مشخص شده از داده به صورت شبكه عصب با گسترش به جلو[۱۶] مي باشند . يكي از انواع ويژه شكه هاي عصبي پرسپترون چندلايه MLP ، كه گراف هاي ساخته شده با فرضيات احتمالي در مورد داده پارامترگذاري مي شوند . شبكه عصبي پياده سازي شده در اين جا شامل نودهاي مخفي و لايه خروجي با در نظر گرفتن يك لايه مخفي مي باشد . روابط ما بين خروجي y و ورودي x به صورت زير تعريف مي گردد :

$$y_k = f_{\text{outer}}\left(\sum_{j=1}^{M} w_{kj}^{(2)} f_{\text{inner}}\left(\sum_{i=1}^{d} w_{ji}^{(1)} x_i + w_{j0}^{(1)}\right) + w_{k0}^{(2)}\right)$$

$w_{ji}^{(1)}$ و $w_{ji}^{(2)}$ وزن ها در لايه اول و لايه دوم به ترتيب با ورودي I و لايه مخفي j مي باشد و $w_{j0}^{(1)}$ باياس براي واحد مخفي j است . M تعداد واحدهاي مخفي ، d تعداد واحد ورودي و k شاخصي براي واحد خروجي مي باشد . تابع $f_{\text{outer}}(\cdot)$ خطي است و $f_{\text{inner}}(\cdot)$ تابع تانژانت هايپربوليك مي باشد . روش بيزين توزيع وزن ها در معادله بالا مي باشد . مساله مشخص كردن وزن هاي شبكه با قرار دادن در فرم بيزين مي باشد .

$$P(w|D) = \frac{P(D|w)P(w)}{P(D)}$$

$P(w)$ تابع توزيع احتمال در فضاي وزني با فقدان داده مي باشد و به صورت تابع توزيع اوليه مي باشد و $D \equiv (y_1, \ldots, y_N)$ ماتريس شامل داده خروجي مي باشد . كميت $P(w|D)$ توزيع احتمال وزن ها با تابع توزيع اختمال بعدي بعد از داده ديده مي شود $P(D|w)$ تابع توزيع احتمال و $P(D)$ تابع توزيع احتمال ثانويه است كه به صورت معادله زير مي باشد :[۱۹]

$$P(w|D) = \frac{1}{Z_s} \exp\left(-\beta \sum_{n}^{N} \sum_{k}^{K} \{t_{nk} - y_{nk}\}^2 - \frac{\alpha}{2} \sum_{j}^{W} w_j^2\right)$$

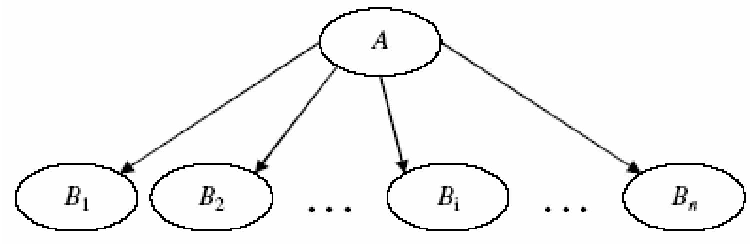
و

$$Z_s(\alpha, \beta) = \int \exp\left(-\beta \sum_{n}^{N} \sum_{k}^{K} \{t_{nk} - y_{nk}\}^2 - \frac{\alpha}{2} \sum_{j}^{W} w_j^2\right) \mathrm{d}w$$
$$= \left(\frac{2\pi}{\beta}\right)^{N/2} + \left(\frac{2\pi}{\alpha}\right)^{W/2}$$

[۱۶] feed-forward

اولين ضابطه در تابع احتمال و دومين ضابطه اطلاعات اوليـه اسـت n شـاخص بـراي الگـوي آموزشـي ، β داده خطادار و k شاخص براي واحد خروجي α ضريب اطلاعات اوليه است . آموزش شبكه بـا اسـتفاده از بيـزين بـه صورت خودكار با پيچيدگي بالا مي باشد . [۱۹]

۲-۶- شبكه هاي بيزين ساده : مدل شبكه هاي بيزين ساده توسط Titterington در سـال ۱۹۸۱ پيـشنهاد شد . شكل زير نمايش مدل شبكه هاي بيزين ساده را نشان مي دهد .



مدلي از شبكه بيزين ساده

در اين مدل از شبكه ها نود مورد نظر (مورد علاقه) به عنوان نود ريشه مي باشد كه اين نود ، نـود والـد ديگـري ندارد . $B_1, B_2, \ldots, B_n$ متغيرهاي ما مي باشند . در اين مدل شروط مستقل از هم به صـورت زيـر در نظـر گرفته مي شوند :[۲۸]

$$B_i \perp \{B_1, B_2, \ldots, B_{i-1}, B_{i+1}, \ldots, B_n\} | A \quad \text{for } i = 1, 2, \ldots, n.$$

مطابق فرضيات شروط مستقل فرض مي كنيم عدم تساوي بعدي A مي تواند به صورت زير توصيف شود

$$\frac{P(A|B)}{P(\sim A|B)} = \frac{P(A)}{P(\sim A)} \times \prod_{i=1}^{n} \frac{P(B_i|A)}{P(B_i|\sim A)}.$$

B نمايش بردار $(B_1, \ldots, B_n)$ مي باشد . اگر و فقط اگر K شامل n پيش گـويي كننـده باشـد آنگـاه عـدم تساوي بعدي براي A مشابه با معادله بالا خواهد شد و فقط از نسبت تشابه K استفاده مي شود . [۲۰]

۲-۷- الگوريتم كاهش داده بيزين: الگوريتم كاهش داده بيزين توسط Lynch & Willett در سالهاي ۱۹۹۹ و ۲۰۰۳ پيشنهاد شد . اين الگوريتم يك روش دسته بندي بر اساس فرضيه احتمال سمبلهاي گسسته در هر كلاس است كه به طور يكنواخت توزيع شده است . انتخاب ويژگيها و ساختار مدل با روش خودكارسازي بالا و تعدادي از تصميمات در مدل ساخته مي شود . در اين الگـوريتم همـه داده هـا نامگـذاري شـده و بـه صـورت علامتـدار يـا سمبليك نمايش داده مي شوند . بنابراين هر سمبل با يك احتمال اوليه فرض مي شود . چرا كه همـه داده هـاي اوليه نمايش داده نمي شوند . پيش پردازش داده الگوريتم كاهش داده بيزين با گسستگي داده بـه شـكل مقـادير قبل از انتخاب اندازه صورت مي پذيرد . براي نمونه اگر چـارك (آمـاري) در سـطح اوليـه انتخـاب شـود و سـپس الگوريتم با متغيرهاي حقيقي مقداردهي شود به چهار متغير در سطح (0,1,2,3) تبديل مي شود . انگيـزه بـراي انتخاب مقادير گسسته برروي مكان اوليه غير آموزنده بر روي بردارهاي ويژگي است كه بر روي مقـادير پيوسـته مثمر ثمر نيست . از اين رو تعداد مساوي نقاط داده در هر چندك (آمارشناسي) مشخص مي شود . مقـادير مـي توانند به صورت يكنواخت توزيع شده و به صورت هدايت پذير توزيع گردند . براي داده ورودي گسسته عموما اين روش سطوح تدريجي اوليه را تحت تاثير قرار نمي دهد و مي تواند با تعداد مساوي دسته بنـدي گسـسته تنظـيم شود . اما براي داده پيوسته فرآيند گسستگي (تمايز پذيري) ممكن است باعـث از بـين رفـتن اطلاعـات گـردد .

سطوح تدریجي [۱۷] در این الگوریتم مي توانند در سطوح خیلي بالا مثلا ۱۰۰ سطح تنظیم گردند . یك مرتبه همه متغیرها به صورت مقادیر اسمي نمایش داده مي شوند الگوریتم براي اصلاح دنباله اي از ویژگیها به صورت معکوس به کار مي رود . جستجوي فرآیند براي کاهش ویژگیهاي نامربوط از داده آموزشي در هر کلاس صورت مي پذیرد . به طور خلاصه الگوریتم ویژگیها را انتخاب کرده و بهترین دسته بندي را انجام مي دهد . **[۲۵]**

۳- کاربرد شبکه هاي بیزین :

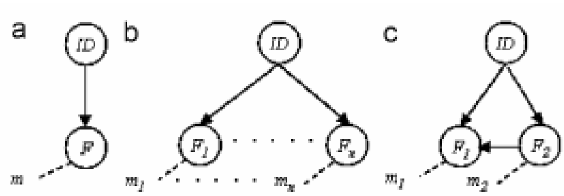
۳-۱- پیوند یا ترکیب [۱۸] چند گانه بیومتریك ها با تخمین کیفیت :

براي افزایش امنیت بالا، بیومتریك چندگانه ترکیبي ، براي کمك به کاهش خطا مورد توجه قرار گرفت . چندین سیستم اثر انگشت براي مثال ، از تمام ترتیب اثرگذار قابل دسترس تا زماني که یك تطبیق قابل قبول بدست بیاید استفاده کردند . اپراتورهاي منطقي دیگر مانند (AND/OR) و جمع توانست با توافق فرد خبره (کارشناس) مورد ملاحظه قرار گیرد . اگرچه روشهاي بیزین در بیومتریك ها مورد استفاده قرار مي گیرند ، قدرت و انعطاف پذیري شبکه هاي BBN [۱۹] تاکنون به طور کامل مورد بهره برداري قرار نگرفته است . ویژگي هاي مشخص به صورت یکپارچه از طریق شکل و توزیع و احتمال مدل گذاري ذاتي به عنوان متغیر شرطي محاسبه مي شوند . این توزیع ها به صورت داده هاي آموزشي آماري که مي تواند به صورت پیوسته ، گسسته و ترکیبي باشد تعریف مي شوند . معماري BBN مي تواند علامت چندگانه فرآیند از یك بیومتریك منفرد یا مجموعه اي از علایم بیومتریك چند گانه در محدوده اي از چارچوب دسته بندي کننده منفرد باشد . **[۲]**

۳-۱-۱- معماري ترکیبي : BBN براي نمایش گرافیکي توزیع احتمال متصل است که به صورت زیر تعریف مي شود .

$$P(X_1, \ldots, X_n) = P(X_n)\prod_{i=1}^{n} P(X_i|\text{parents}(X_i)),$$

$\text{parents}(X_i)$ متغیر تصادفي X_i در شرط مي باشد . شبکه شامل مجموعهرئوس یا نودها متناظر با متغیرهاي تصادفي و یك مجموعه از لبه هاي جهت دار مابین جفت هاي راس از یك گراف جهت دار غیر چرخشي است که در لبه هاي جهت دار به شکل نود متناظر با X_i مي باشد که به عنوان والد X_i تعریف مي شود . نودهاي بدون والد نود ریشه نامیده مي شوند . تعریف با توزیع شرطي $P(X_i|\text{parents}(X_i))$ و چگالي احتمال اولیه براي نودهاي ریشه تعریف مي گردد . از این رو فاکتورهاي در معادله منحصر به فرد نیستند . روابط ما بین متغیرهاي $X_1, \ldots, X_n$ مي تواند با گراف جهت دار غیر حلقوي و مجموعه اي از احتمالات شرطي $P(X_i|\text{parents}(X_i))$ با توزیع منحصر به فرد $P(X_1, \ldots, X_n)$ در معادله بالا تعریف شود . شکل زیر شبکه هاي بیزین ساده BBN براي تشخیص هویت را نشان مي دهد .



متغیر F براي تطبیق بیومتریك چند گانه و محاسبه مابین الگوهاي منحصر به فرد و مقدار مستقل بر روي حالت متغیر هویت ID نمایش داده مي شود . مانند : میزان تطبیق ما بین دو نمونه از $(ID = 1)$ منحصر به فرد یکسان که اگر از نمونه هاي منحصر به فرد $(ID = 0)$ متفاوت بودند باید بزرگتر شوند . براي دسته بندي کننده در شکل a

[۱۷] quantization
[۱۸] Fusion
[۱۹] Baesian belief network

معادلـه $P(X_1, \ldots, X_n) = P(X_n) \prod_{i=1}^{n} P(X_i|\text{parents}(X_i)),$ باعـث کـاهش $P(F, ID) = P(F|ID) \cdot P(ID)$ مـي شود . ميزان تطبيق m توزيع بروز شده هويت به صورت زير مي باشد :

$$P_{update}(ID) = P_{prior}(ID) \left[\frac{P(F=m|ID)}{\sum_{ID} P(F=m|ID)} \right].$$

از اين رو :

$$\frac{P_{update}(ID=1)}{P_{update}(ID=0)} = \frac{P_{prior}(ID=1)}{P_{prior}(ID=0)} \times L,$$

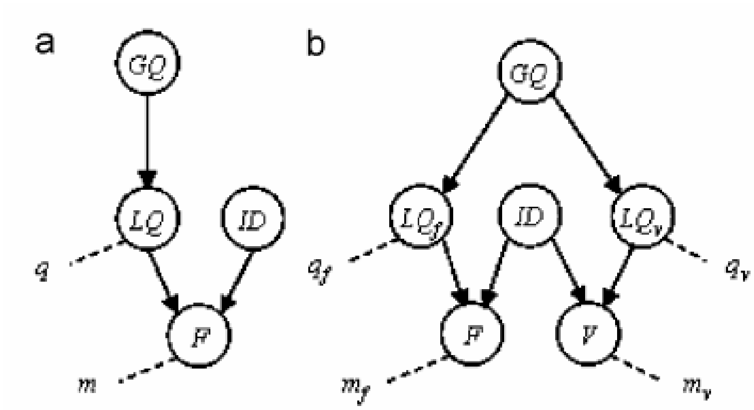
در اين صورت $L = P(F=m|ID=1)/P(F=m|ID=0)$ نرخ احتمال مي باشد . نـرخ احتمـال بـه صـورت آمـاري تست شده و در بيومتريك مورد استفاده قرار مي گيرد. توزيع اوليه نامساوي بـه صـورت يكپارچـه مـي باشـد و اگـر شناخته شود به عنوان فاكتور وزني مي باشد . براي مثال در هر روز دسترسي به عمليات كنترل صورت مي گيـرد و امكان دارد به صورت غير محتمل به عنوان ورودي اوليه در نظر گرفته شـود و بـا مجموعـه اي از شـرايط اوليـه بـه صورت $P_{prior}(ID=0) \ll P_{prior}(ID=1)$ باشد . دسته بندي كننده شكل b مستقل از معادله بالا اسـت و به صورت زير مي باشد :

$$L = \prod_{i=1}^{n} \frac{P(F_i = m_i|ID=1)}{P(F_i = m_i|ID=0)}.$$

براي جزييات بيشتر مي توانيد به اين مقاله مراجعه نماييد . **[۲]**

۳-۱-۲- به كارگيري كيفيت در تركيب بيومتريك ها : كيفيت مي تواند معاني مختلفي داشته باشـد كـه تحت وضعيت هاي مختلف صورت مي پذيرد . كيفيت معيار قابل سنجش (اندازه گيري) است كه به دقت و صحت و نرخ سيگنال نويزي دلالت مي كند . اگرچه قابليت اعتماد بر روي دنباله قابل سنجش و يا قابليت اعتماد به عنوان تشخيص دهنده قابل مشاهده مي باشد . ميزان مشخص كژريختي (كج و كولگي شكل) با جزييات كمتر از اثر انگشت داراي نويز يا خطاهايي مانند اغتـشاش محلـي بـر روي پوست انعطاف پذير قابل اشتقاق مي باشد . اگرچه الگوريتم هاي تطبيق بـراي ايـن ناپايـداري (عـدم قطعيت) طراحي شده اند . پاك كردن چندين ناحيه مخفي و بي نظم مانند عبور از راههاي اشتباه و ادغام پلها ، مي تواند تاريك و مبهم باشد . معرفي ساختگي دو راهـي هـا ۰ (محـل انـشعابات بـه دو بخش) قابل معرفي است . امكان دارد تطبيق به صورت مخالف با معرفي جزييات كم اهميـت كـاذب تحت تاثير قرار گيرد . عموما ميزان تطبيق از اثر كيفيت ضعيف كه قابل اطمينان نمي باشد ، اشتقاق مي شود . به هر حال اگر ميزان تطبيق بالا بود از يك ناحيه داراي خطا [۲۰] شدني نيـست . بنـابراين ميزان كيفيت پايين مي تواند اطلاعات هويت را مشخص كرده و نبايد فورا رد شود . **[۲]**

[۲۰] imposter



BBNs براي تركيب اثرانگشت و صدا . نودهاي **V** و **F** متناظر با ميزان تطبيق صدا و اثر انگشت

شكل a و b مدلهاي BBN براي اثر انگشت و خصوصيات اثر انگشت به علاوه صدا مي باشند . كيفيت نودها با LQ نمايش داده مي شود . كيفيت به صورت مشخص ميزان تطبيق را تحت تاثيز قرار مي دهد . اما هيچ رابطه اي ما بين كيفيت و هويت وجود ندارد و اين استقلال در گراف منعكس شده است . ويژگي جديد در اين مقاله متغير كيفيت سراسري GQ مي باشد كه به عنوان معيار اندازه گيري اطمينان كه همبسته به $P_{update}(ID)$ براي تصميم گيري در نظر گرفته مي شود . توزيع شرطي $P(LQ|GQ)$ يك بخش از تعريف GQ است كه براي داده آموزشي مورد ارزيابي و تخمين قرار نمي گيرد . ما مي خواهيم GQ و LQ به صورت مثبت به همديگر همبسته باشند . بنابراين به محورهاي در يك حوزه با چگالي احتمال پيوسته $P(LQ, GQ)$ در يك خط GQ=LQ نياز داريم . از اين رو فض مي كنيم متغيرهاي كيفي سراسري و محلي شامل مقدار دامنه و گستردگي يكساني باشند . مي توان گفت $\Theta=\{1, 2, \ldots, K\}$ توزيع متصل (پيوسته) اي است كه به صورت شبكه $\Theta \times \Theta$ تعريف مي شود :

$$P(LQ=q_1, GQ=q_2) := \frac{f(q_1-q_2)}{\sum_{(q_1',q_2')\in\Theta} f(q_1'-q_2')}$$

براي كليه $(q_1, q_2) \in \Theta$ ، f يك تابع متقارن از مبدا مي باشد . بنابراين خواهيم داشت :

$$\Omega := \sum_{(q_1,q_2)\in\Theta} f(q_1-q_2) = Kf(0) + 2[(K-1)f(1) + 3(K-2)f(2) + \cdots + f(K-1)].$$

Then

$$P(LQ|GQ=q) = \frac{1}{\Omega} f(LQ-q).$$

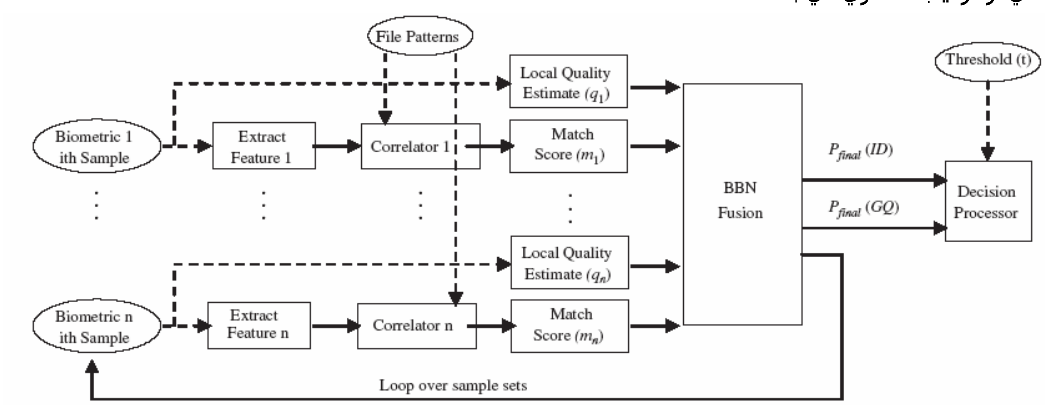
براي تعاريف اوليه $f(x)$ را به صورت ميانگين گوسي صفر با $\sigma=2.8$ در نظر مي گيريم كه LQ وGQ مرز نمي باشند . حال مي توان اين تعريف را اصلاح نمود . بروز رساني معادله براي شكل در حالت b به صورت

$$
\begin{aligned}
&\frac{P_{update}(ID)}{P_{prior}(ID)} \\
&\quad = P(F = m_f | ID, LQ_f = q_f) P(V = m_v | ID, LQ_v = q_v) \\
&\qquad \times \sum_{GQ=1}^{5} [P(LQ_f = q_f | GQ) \\
&\qquad \times P(LQ_v = q_v | GQ) P_{prior}(GQ)]
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{P_{update}(GQ)}{P_{prior}(GQ)} \\
&\quad = P(LQ_f = q_f | GQ) P(LQ_v = q_v | GQ) \\
&\qquad \times \sum_{ID=0}^{1} [P(V = m_v | ID, LQ_v = q_v) \\
&\qquad \times P(F = m_f | ID, LQ_f = q_f) P_{prior}(ID)].
\end{aligned}
$$

مي باشد كه مشابه معادله در حالت a است معادله مي تواند شامل ويژگيهاي اضافي به صورت توسعه يافتـه باشـد . متغيرهاي شرطي و مستقل مانند انواع انگشت ، سن ، جنسيت ، قوميت و غيره مي باشد . اگر داده هـاي بـه انـدازه كافي و حوزه دانش وجود داشت چگونگي تمايز اين تاثير در ميزان كيفيت توزيع نـشان داده خواهـد شـد . نمـايش گرافيكي به صورت روابط شرطي مابين متغيرهاي مربوط قابل فهميدن صورت مـي پـذيرد . بـا افـزايش پيچيـدگي بروزرساني معادله ممكن است از راه تجربه طولاني تر نباشد . فوايد ديگر نمايش گرافيكي منجر به الگـوريتم بهينـه سازي براي كاربرد قوانين بيز است كه مي تواند به طور مستدل باعث كاهش محاسبات شود . شكل زير يك معماري كلي از تركيب معماري مي باشد .

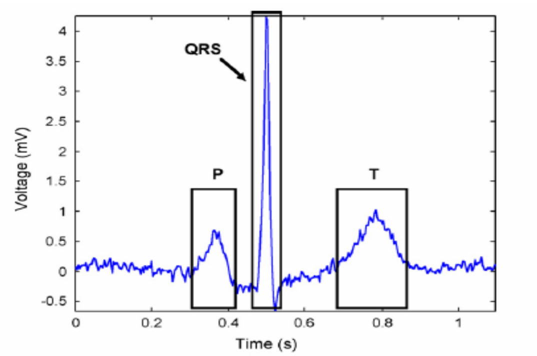


معماري پيشنهاد شده **BBN** است كه مبتني بر خصوصيات معماري تركيبي نمونه ها و كيفيت بيومتريك ارزيابي شده در **BNN** براي توليد احتمال چگالي و متغيرهاي تصادفي چگالي كه $ID = 1$ براي هويت و $ID = 0$ براي ساير موارد تركيب شده مي باشد . و **GQ** متغير تصادفي از كيفيت سراسري است .

شكل بالا مي تواند سازگاري براي ورودي هاي شامل نمونه ها ي چندگانه از بيومتريك هاي يكـسان يـا بيومتريـك هاي چند گانه باشد . كه مي تواند به طور همزمان فرآيند پردازش ترتيبي يك مجموعه قابل اندازه گيـري و مقـادير وابسته ما بين آنها باشد . BBN دقتي به ويژگي ندارد و الگوريتم هاي تطبيق را مورد اسـتفاده قـرار داده و ورودي آن ميزان و اندازه ويژگي به عنوان رفتار آماري به صورت احتمال شرطي در شبكه را نمـايش مـي دهـد . نرمـاليزه

كردن به صورت غير صريح در توزيع شرطي استفاده شده و كيفيت هاي مختلف در ويژگيهـاي طبيعـي بـه عنـوان متغيرهاي شرطي تحت تاثير قرار مي گيرند . [۲]

۳-۲- دسته بندي كننده بيزين براي دستگاه نوار قلب مبتني بر سن افراد در كاربردهاي پزشكي : يك روش براي دسته بندي كردن بيماران بر اساس ويژگيهاي آماري استخراج شده از سيگنالهاي قلـب اسـت كه به طور معمول در دسته بندي كننـده شـبكه بيـزين تكامـل يافتـه مـورد اسـتفاده قـرار مـي گيـرد . متغيرهاي سيگنال پيوسته به سمبل هاي گسسته توسط آستانه به سيگنالهاي با ابعاد پايين تر تبديل مي شوند . اين محاسبات ساده جدول احتمال شرطي است كه براي دسته بندي كننده اسـت . ابتـدا از روش جستجوي قله نوردي به صورت حريصانه و دوم از الگوريتم ژنتيك براي محاسـبات تكـاملي اسـتفاده مـي شود . قلب انسان يك سيستم پيچيده اي است كه دستگاه نوار قلب الهام گرفته از آن مي باشد . آمـوزش فيزيكي قادر به تشخيص الگوها در سيگنالهاي نوار قلبي از بيماران و براي تشخيص بيماري مورد استفاده قرار مي گيرد . [۳]



سيگنالهاي نوار قلب با **QRS** پيچيده با**P** و **T** امواج تشخيص داده شده

دسته بندي كننده شبكه بيزين به تشخيص ما بين دو گروه منحصر بـه فـرد بـر اسـاس ويژگيهـاي مـشتق شـده از دستگاه نوار قلب [۲۱] بر اساس سلامتي افراد جوان و ميانسال مي پـردازد . اگـر ايـن روش بـه تـشخيص دو گـروه از بيماران پرداخت ، دسته بندي كننده پيچيده تري در مساله مانند طبقه بنـدي احتمـال خطـر بيمـاران قلبـي مـي پردازد كه براي اندازه گيري باليني تركيب چند گانه مي باشد و قادر به دقت بيشتر از روش هاي سنتي مـي باشـد ، مورد استفاده قرار مي گيرد . [۳]

۳-۲-۱- پردازش سيگنال دستگاه نوار قلب : براي تشخيص مشخـصات رفتـاري در دسـتگاه نـوار قلبـي ، تعداد كيفيت هاي مورد نياز براي تجزيه سيگنال مولفه هايي است كه مي تواند تجزيه و تحليل شده و به صورت كامپيوتري مورد استفاده قرار گيرد . بسياري از انواع ويژگيهاي (خاصـيت ، كيفيـت) [۲۲] استخراج شده و به عنوان ويژگيهاي هاي سيگنال در اينجا مراجعه مي كنند كـه اغلـب تركيـب آنهـا مشكل است و بايد به صورت هدفمند ارزيابي شده و راهي مفيد براي كاربرد باشد . براي نمونـه يـك الگوريتم براي تشخيص QRS در دستگاه نوار قلب گزارش شده پيچيده است و يك چـارچوب بـراي دسته بندي كننده ۱۲ شاخص دستگاه نوار قلب به صورت تعاملي براي تجزيه و تخليل سـيگنالها بـه كار مي رود . امروزه دو روش اساسي براي پردازش سيگنال به صـورت بيولـوژي وجـود دارد . اولـين روش متغير پنجره اي [۲۳] است كه شامل قطعه بندي داده با طولهاي كوچك مـي باشـد. ايـن قطعـه بندي به طور يكنواخت يك فضاي يكنواخت است كه با سيگنال مبـدا و ممكـن اسـت بـا نقـاطي در

[۲۱] electrocardiograms (ECG)
[۲۲] quantifications
[۲۳] sliding window

پنجره هاي ديگر همپوشاني نداشته باشد . يك ويژگي يا معيار آماري براي هـر قطعـه محاسـبه مـي شود . مانند مقدار فركانسي يا انحراف استاندارد دامنه سيگنال كه هر كـدام از ايـن محاسـبات مـي توانند به صورت موقتي مطابق موقعيت سيگنال مبدا به صورت پيشرفت ويژگيها در سراسر زمـان در كل پنجره نمايش داده شوند . سيگنال فشرده شده ، نمايش داده مي شود و تـا زمـاني كـه جـستجو براي تغيير حالت رخداد فيزيكي منجر مي شود ، مفيد مي باشد . مانند تعريف شروع امواج مغزي يـا حمله قلبي از ECG . زماني كه سيگنال مهم نيست يا تغيير حالت انتظار مي رود روش متغير پنجره چندان پيشنهاد نمي شود . به همبن علت دومين روش محاسبه ويژگي سيگنال بيولوژيكي است كه براي محاسبه ويژگي بر روي كل سيگنال محاسبه مي گردد. ويژگيهاي زيادي همانند تبـديل و آمـار مي تواند در سيگنال دستگاه نوار قلبي به كار رود ، كه شامل امواج ، فركانس و تبديلهاي غير خطـي به صورت آماري ، موقتي و ويژگيهاي بي نظم استخراج شده است كه مي تواند به صورت پـشت سـر هم اجرا شود . زماني كه الگوهاي مخفي يا انحراف هاي نامتراكم در شكل سيگنال يا ويژگيهاي ديگر كشف مي شوند اهداف اصلي كه فشردگي ، كدگذاري يـا نقـاط مـشخص كـشف شـده در سـيگنال دستگاه نوار قلب است را شامل مي شود . **[۳]**

۳-۲-۲- استخراج ويژگي سيگنال در **ECG** : سيگنالهاي ۱۲۰ دقيقه اي براي محاسبه ويژگيهاي سيگنال و محاسبه ۱۲ ويژگي استخراج شده در شكل زير به كار مي روند . **[۳]**

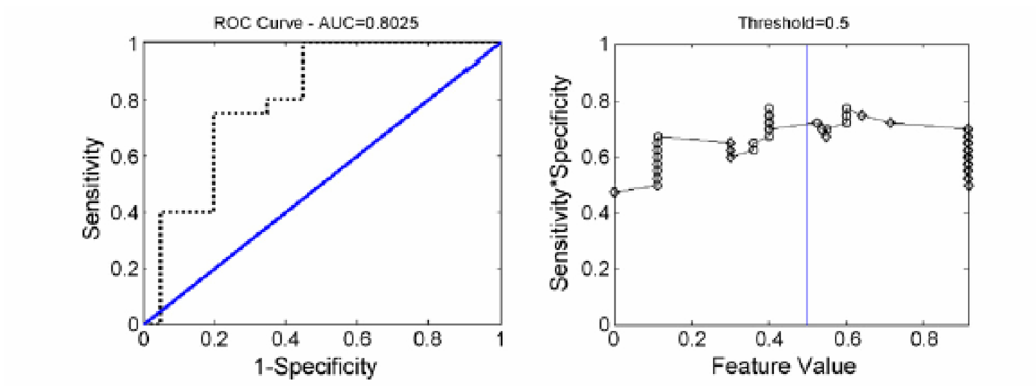
بلوك دياگرام از ويژگيهاي استخراج شده

$$Energy = \sum x_i^2$$
$$4^{th}\ Power = \sum x_i^4$$
$$Nonlinear\ Energy = \sum -x_i \cdot x_{i-2} + x_{i-1}^2$$
$$Curve\ Length = \sum x_i - x_{i-1}$$
$$Hurst = \ln\left(\frac{range(x_i)}{std(x_i)} - \frac{i}{2}\right)$$
$$Peak\ Power = \max(PSD)$$
$$Peak\ Frequency = index(\max(PSD))$$
$$Mean\ Frequency = index(mean(PSD))$$
$$Median\ Frequency = index(median(PSD))$$
$$Spectral\ Entropy = \sum PSD \cdot \log(PSD)$$
$$Shannon\ Entropy = -\sum hist(x) \cdot \log(hist(x))$$
$$Katz\ Fractal\ Dimension = \sum_{i=1}^{k} \frac{\log(k-1)}{\log\left(\frac{\max(\sum_i \sqrt{(x_i - x_1)^2 + i^2})}{\sum_i \sqrt{(x_{i+1} - x_i)^2 + 1}}\right) + \log(k-1)}$$

معادله اي از ويژگيها در **F** ، و محاسبه بر روي سيگنالهاي **ECG**

دليل اصلي براي انتخاب اين ويژگيها بر اساس علايم و زمينه در پردازش سيگنال است . انرژي و تـوان چهـارم مـي تواند به صورت سكون سيگنال هدف در بالا و پايين خط هدف نمايش داده شود . آنتروپي شانون معياري اسـت كـه به صورت تصادفي مقادير نوسان دامنه سيگنال را اندازه مي گيرد . آنتروپي بـالاتر نـامنظم تـر اسـت وبـه سـيگنال تصادفي در گردش برون [۲۴] نزديك تر است . حوزه فركانسي مبتني بر ويژگيها مي تواند مشخصات مهم در سيگنال را آشكار نمايد . پيك فركانسي و توان پيك در سيگنال فركانسي آشكار است . بعد از استخراج مقـادير ويژگيهـا بـه شكل باينري مبتني بر مقادير در بالا و پايين آستانه مشخص تميز داده مي شوند . اين آستانه مجموعـه ايـست كـه درمنحني مشخصات عملياتي گيرنده [۲۵] به كار مي رود مقادير ويژگي مانند سن قابل پيشگويي مي باشند .

[۲۴] Brownian
[۲۵] Receiver Operating Characteristic (ROC)



منحني **ROC** – نمايش مربعي بيماران با گروه سني ميانسال است .و نمايش دايره اي گروه سني جوان را نشان مي دهد خط عمودي به تعريف مقادير آستانه براي نمايش ويژگي و فركانس مي پردازد .

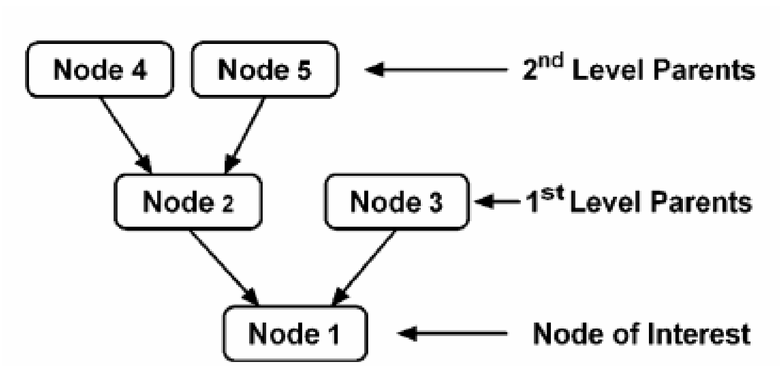
۳-۲-۳- آشكار سازي ساختار شبكه بيزين : احتمالات شرطي در مفاهيم پزشكي بـه كـار مـي رونـد . C علامت پزشكي و E روابط مستقل موجود ما بين E و C مي باشد . اين احتمال به صورت زير نمايش داده مي شود . **[۳]**

$$P(C|E) = \frac{P(E|C) \cdot P(C)}{P(E)}$$

تئوري بيز روشي است كه براي پيدا كردن احتمال معكوس شرطي به كار مي رود :

$$P(E|C) = \frac{P(C|E) \cdot P(E)}{P(C)} = \frac{P(C,E)}{P(C)}$$

مساله با متغيرهاي باينري n قابل ملاحظه است كه روابط ما بين آنها براي پيشگويي متغير خروجي كلاس سـيگنال شناخته شده نيست . مانند نود ۱ در شكل زير :



ساختار شبكه بيزين و ترمونولوژي

اگر تمامي متغيرهاي مرتبط در توزيع پيوسته منفرد به كار روند تمامي نودها در اولين سطح ازوالـدين هـم ارز مـي باشند و تعداد تركيب هاي ممكن از متغيرها برابر با $(2^n - 1)$. مي شود . براي هر تركيب تعداد كافي از نمونه ها رخ دهد و يك تخمين احتمال انعطاف پذير (كشساني) بدست آيد . اگر روابط مـستقل مـا بـين متغيرهـا توانـست تعريف شود نتايج در متغيرهاي مستقل حذف مي شود (برداشته مي شود) . نودهاي جزيـي مجـاور نودهـاي مـورد نظرند . اين نود حذف شده منجر به كاهش تعداد متغيرهاي تركيبي مي شود بنابراين مقدار عامـل كاهنـده بـه داده نياز دارد . متغيرها شروط جهت دار مي باشند و جز نودهاي مورد نظر نمي باشند بنابراين والدين نودها مورد علاقه (مورد نظر) مي باشند . (مانند نودهاي ۴ و ۵ در نسبت با نود ۱) . و براي سيستم هاي بزرگ زماني كه تعـداد داده از بين مي رود توزيع مي شود . و زماني كه نقاط داده از دست رفته رخ داد استنتاج مبتني بر بيزين به كار مي رود . توسط متغيرها احتمال شرطي تعريف مي شود و مقادير احتمالي حذف شده شامل نقاط از دست رفته مـي باشـند .

matlab1.ir

شبكه بيزين قادر به دستكاري نقاط داده از دست رفته و مستلزم پاين آوردن اطلاعـات مبتنـي بـر دانـش اوليـه در سيستم ها با متغيرهاي مستقل با فوايد مهم است . در اين روش دو شبكه قابل مقايسه است اولين روش جستجوي قله نوردي به صورت حريصانه و دومين روش محاسبه با استفاده از الگوريتم ژنتيك است . كـارايي ايـن دو روش بـا صحت بدست آمده از دسته بندي كننده بيزين قابل مقايسه و اندازه گيري مي باشد .[۳]

۳-۲-۴- دسته بندي كننده ساده : يـك مجموعـه از علايـم $E = \{E_1,\ E_2,\ \ldots,\ E_n\}$ را داريـم كـه احتمال پيوسته (متصل) در معادلـه $P(E|C) = \frac{P(C|E) \cdot P(E)}{P(C)} = \frac{P(C,E)}{P(C)}$ شـماره گذاري شده است . $P(C,E)$ با استفاده از تعريف احتمال شرطي به كار برده شده توسعه يابد :

$$\begin{aligned} P(C,E_1,\ldots,E_n) &= P(C) \cdot P(E_1,\ldots,E_n|C) \\ &= P(C) \cdot P(E_1|C) \cdot P(E_2,\ldots,E_n|C,E_1) \\ &= P(C) \cdot P(E_1|C) \cdot P(E_2|C,E_1) \cdot P(E_3,\ldots,E_n|C,E_1,E_2) \\ &= P(C) \cdot P(E_1|C) \cdot P(E_2|C,E_1) \cdot P(E_3|C,E_1,E_2) \\ &\quad \cdot P(E_4,\ldots,E_n|C,E_1,E_2,E_3) \end{aligned} \quad ($$

اين توسعه مستلزم به كارگيري علايم مهم در تعريف داده در بسياري از احتمالات اسـت . بـه همـين منظـور بـراي كاهش تعدادي نمونه هاي داده لازم داريم . يك فرض مستقل براي مولفه E ساخته مي شـود احتمـال پيوسـته بـه صورت زير است :

$$\begin{aligned} P(C,E_1,\ldots,E_n) &= P(C) \cdot P(E_1|C) \cdot P(E_2|C) \cdot P(E_3|C) \cdots \\ &= P(C) \prod_{i=1}^{n} P(E_i|C) \end{aligned}$$

براي فرضيات مستقل از دسته بندي كننده (ساده)[۲۶] استفاده مي نماييم . شبكه ايجاد شده فقط با يك نـود بـراي نمايش تمام نودها در اولين سطح والد به كار مي رود. . [۳]

۳-۳- تشخيص سرطان سينه با استفاده از شبكه بيزين : دسته بندي (طبقه بندي) به انتـساب برچـسب ها به كلاس ها براي نمونه هاي بدون برچسب گذاري دلالـت مـي كنـد . يـك مجموعـه از نمونـه هـاي برچسب گذاري نشده به صورت دستي و يك سري از آنها راه حل مساله مي باشد كه براي يـافتن تـابع مناسب و نگاشت با نمونه هاي برچسب گذاري نشده متناظر با برچسب آن كلاس به كار مـي رود . كـه مي توان با جستجوي نواحي مورد نظر دسته بندي خودكار طراحي كرد كه كه براي تخمين ايـن تـابع از داده براي نمونه از شبكه بيزين استفاده نمود . اين نوع از يادگيري ، يادگيري بانـاظر اسـت و بـه همـين علت مختصر بوده و فاقد فضا مي باشد . از ۷ زير روال در اين آزمايش استفاده مي نماييم . **a**) - دسته بندي كننده شبكه بيزين ساده **b**) – بيزين –N **c**)- بيزين MP **d**) – حريصانه **e**)- بيزين + MP حريصانه **f**)- CBL2 و **g**)- PC .

a) - يك شبكه بيزين ساده است كه يكي از موثرترين دسته بندي كننده ها مي باشد و در مقابل وضعيت دسـته بندي كننده هاي مصنوعي (ساخته بشر) مقايسه مي شود . جذابيت اصلي اين الگـوريتم سـادگي و دقـت آن مـي باشد اين ساختار هميشه ثابت است (متغير كلاس يك كمان اشاره گر در هر ويژگي است) كـه ايـن دسـته بنـدي كننده شامل دقت دسته بندي كننده بالا و خطاي بيزين بهينه است . ضوابط ساده يادگيري NB از يك نمونه داده

[۲۶] Naive

matlab1.ir

آموزشي و احتمال شرطي در هر كلاس از ويژگي ها مي باشد . يكي از شرط هاي وارد شده NB استفاده از قـوانين بيز براي محاسبه احتمال شرطي مجموعه ويژگي هاي منتخب شده از كلاس با احتمال قبلي بالا مي باشد .

b)- بيزين **N** : بيزين N متعلق به خانواده ي الگوريتم هاي مبتني بر قيود كـه بـراي آزمـايش شـرايط مـستقل استفاده مي شود . اگر يك جفت از نودها متصل يا نامتصل باشند ، بر اساس اندازه گيري اطلاعات تصميم گيـري مي شوند

c)- بيزين **MP** : متعلق به خانواده بيزين N مي باشد با اين تفاوت كه مستلزم ويژگيهاي صريح براي متغيرهـاي كلاسي نيستند . و گرايش به نمايش توزيع احتمالي پيوسته با تعداد حداقل از كمانها دارد .

d)- حريصانه : اين الگوريتم به صورت جستجوي قله نوردي حريصانه براي بخش تبديل به كار مـي رود . در ايـن جا براي آزمايشات گزارش شده زير روال حريصانه به عنوان ورودي يك گراف خالي و يك پايگاه داده مي باشد . بـه طور تناوبي براي فضاي اوليه يك گراف كامل مي باشد . يك گراف تصادفي يا يك گراف بر اساس قيود پيشنهاد داده مي شود .

e)- تركيب الگوريتم حريصانه با بيزين MP است .

f)- نرم افزار شامل دو الگوريتم متفاوت است الگوريتم CBL1 و CBL2 با ساختار شبكه بيزين از داده مي باشد

g)- زير روال ديگر براي ساختار شبكه بزين از داده PC ناميده مي شود .

پايگاه داده شامل ۶۹۲ داده است كه به دو زير مجموعه تقسيم مي شود . كه به طور تصادف اين داده ها به دو نمونه مشترك انحصاري جز بندي مي شوند . داده هاي آموزشي و تست . اندازه مجموعه آموزشـي $\frac{2}{3}$ از داده هـا ۴۶۲ و ما بقي $\frac{1}{3}$ از داده ها ۲۳۰ از مجموعه تست را شامل مي شود . تمامي الگوريتم هاي توصـيف شـده در داده هـاي آموزشي براي يادگيري دسته بندي مورد استفاده قرار مي گيرند و مجموعه تست دقت دسـته بنـدي كننـده ماننـد تعداد زياد نمونه هاي صحيح متناظر با دسته بندي كننده را پيشگويي مي نمايد . براي مقايسه كارايي اين ۷ دسـته بندي كننده حساسيت (توانايي تشخيص صحيح بيماري كه واقعا بيمار است) ، تشخيص (توانايي صحيح كه واقعا بيمار نيست) ، پيشگويي مقادير مثبت حاصل از نتايج (تشخيص احتمال بيماري با نتايج مثبت حاصل از سـرطان و پيشگويي مقادير منفي از نتايج (تشخيص احتمال بيماري با نتايج منفي كه فرد سرطان ندارد) را شامل مي شود . شكل زير و جدول زير نتيجه آزمايشات را نشان مي دهد . **[۵]**

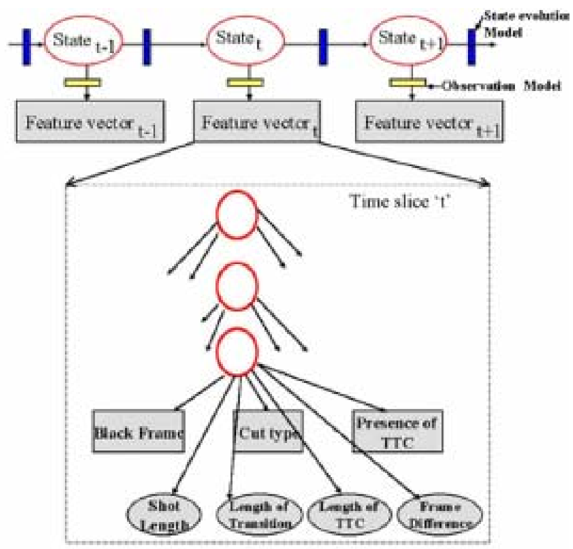
Naive Bayes	Bayes - N	MP-Bayes	Greedy	MP-Bayes + Greedy	CBL2	PC
Outcome⇒Age	TriD⇒Nucleoli	Size⇒Outcome	Age⇒Outcome	Nucleoli⇒Necro	Outcome⇒Age	Age⇒Naked
Outcome⇒Dys	Size⇒Nucleoli	ICL⇒Outcome	Outcome⇒ICL	Nucleoli⇒Dys	Outcome⇒ICL	Outcome⇒Age
Outcome⇒ICL	Nucleoli⇒Outcome	Pleo⇒Size	Outcome⇒TriD	Nucleoli⇒TriD	Outcome⇒TriD	Outcome⇒Dys
Outcome⇒TriD	Pleo⇒Size	Foamy⇒Necro	Age⇒TriD	Size⇒Nucleoli	Naked⇒Foamy	Nucleoli⇒TriD
	Size⇒Outcome	Naked⇒Foamy	Age⇒Naked	Outcome⇒Nucleoli	Necro⇒Foamy	Outcome⇒ICL
Outcome⇒Naked	Necro⇒Nucleoli	Outcome⇒Nucleoli	Age⇒Pleo	Size⇒Outcome	Size⇒Nucleoli	Naked⇒Foamy
Outcome⇒Foamy	Necro⇒TriD	Size⇒Nucleoli	Age⇒Size	Pleo⇒Size	Nucleoli⇒Outcome	Necro⇒Nucleoli
Outcome⇒Nucleoli	Necro⇒Dys	Outcome⇒Age	Age⇒ICL	Pleo⇒ICL	Pleo⇒Size	Nucleoli⇒Outcome
Outcome⇒Pleo	Age⇒ICL	ICL⇒Age	Dys⇒Age	ICL⇒Dys	Outcome⇒Dys	Outcome⇒TriD
Outcome⇒Size	Age⇒Outcome		Age⇒Nucleoli	Dys⇒Necro	Size⇒Outcome	Necro⇒Foamy
Outcome⇒Necro	ICL⇒TriD		TriD⇒Apo	Age⇒Dys	Apo⇒Nucleoli	Nucleoli⇒Size
Outcome⇒Apo	ICL⇒Outcome		Naked⇒Foamy	Age⇒Naked	Naked⇒Age	Apo⇒Nucleoli
			Pleo⇒Nucleoli	Age⇒TriD	Size⇒Age	Size⇒Pleo
			Necro⇒Dys	Outcome⇒Age		Size⇒Outcome
			Necro⇒Foamy	Outcome⇒ICL		Apo⇒Outcome
			TriD⇒Pleo	Naked⇒Foamy		
			TriD⇒Size	Necro⇒Foamy		
			ICL⇒TriD	ICL⇒TriD		
			TriD⇒Nucleoli	TriD⇒Apo		
			Nucleoli⇒Size	TriD⇒Necro		
			Necro⇒Outcome			
			Dys⇒ICL			
			Dys⇒Nucleoli			
			Pleo⇒Size			
			Dys⇒Pleo			
			Dys⇒Outcome			
			Outcome⇒Pleo			

نتايج حاصل از اجراي ۷ دسته بندي كننده بيزين با استفاده از مجموعه داده منفرد به عنوان مجموعه آموزشي

۳-۴- شبکه بيزين موقتي براي کاوش اطلاعات ساختار يافته : داده ويديويي شامل ساختار موقتي قوي است که در فرآيند دسته بندي به کار مي رود . در اين جا به استخراج ساختار موقتي از مشخصات دنباله داده ويديويي به شکل کلاس هاي مختلف مي پردازيم . ويژگي هاي ادراکي شامل زمان تصادم ، طول عکس (فيلمبرداري شده) ، انتقال و فعاليت حرکت موقتي مي باشد . با اين ويژگي هاي ادراکي چندين کلاس ويديويي مشخص شده و منجر به تکوين دنباله اي از دسته بندي سطح بالا مي شود . نتايج پرس و جوهاي سطح بالا به آساني با ويژگيهاي ادراکي نگاشت مي شوند و بهتر از سيستم هاي بازيابي مبتني بر محتوا قابل دسترس مي باشند . ترکيب موقتي ويژگيهاي ادراکي در ساختار سطح بالا مي تواند با شبکه بيزين پويا به کار رود . اين شبکه ها توان استنتاج آماري را اجازه مي دهند و يادگيري با دانش موقتي و بافتاري در مساله ترکيب مي شود . يکي از اهداف سيستم هاي بازيابي مبتني بر محتوا قابليت دسترسي آن براي کاربر مي باشد . از طرفي پرس و جو در ضوابط ويژگي هاي سطح بالا جمله بندي مي شود يا پرمعني تر از سيستم هاي سطوح نحوي نشان گذاري مي شوند . مثالهايي از پرس و جوهاي سطح بالا : توصيف فعاليت آبجکت منفرد ، توصيف فعل و انفعالات آبجکت هاي چندگانه و وقايع بازشناسي ، دسته بندي در انواع فيلم مانند پرتاب تنيس ، بسکتبال و غيره و دسته بندي در ضوابط ناقص (آسيب ديده) و متحرک مي باشد .

۳-۴-۱- شبکه هاي بيزين پويا : Yu and Wolf مدل مارکف مخفي تك بعدي براي صحنه هاي دسته بندي ورودي و خروجي با استفاده از الگوهاي يادگيري آماري در سال ۱۹۹۵ نمايش دادند . Yang and Kuo (1999) در سال ۱۹۹۹ زير روال سلسله مراتبي با استفاده از دو سطح دسته بندي بر اساس دسته بندي کننده kهمسايه نزديك (مجاور) پيشنهاد کردند . دسته بندي نا دقيق در ويژگيهاي تصاوير سطح پايين استفاده مي شود . و دسته بندي سطح خوب مبتني بر معناشناسي مي باشد . Naphade et al. در سال ۱۹۹۸ چارچوب مارکف براي ساخت احتمال اشيا چند رسانه اي که چند کليدي [۲۷] ناميده مي شود و براي ترکيب ويژگيهاي سطح پايين خصوصيات چند گانه به کار مي روند . احتمالات براي رمزگذاري روابط سطح بالا ما بين چند کليدي ها ، بهسازي يا کاهش احتمالات همزمان در چند کليدي هاي مختلف به کار مي رود . زماني که دسته بندي به وضوح مشخص مي شود اين کار نشان داده مي شود . و به علاوه نمايش حوزه اي از داده هاي آموزشي بزرگ و يادگيري در ساختار مدل به کار مي رود . DBN[۲۸] يك کلاس از مدل احتمال گرافيکي است که براي مجموعه اي از متغيرهاي تصادفي مستقل در زمان کدگذاري مي شود . که آنها مدلهاي مارکف مخفي و سيستم هاي پوياي خطي با پذيرش محدوده وسيع از توپولوژي و الگوريتم هاي استنتاج توليد مي شوند .شبکه بيزين پويا از ترکيب ويژگيهاي متنوع به صورت موقتي مانند کشف کننده صورت ، پوست ، سکوت و تشخيص صوت (از طريق صحبت کردن) استفاده مي نمايند . کاربردهاي مهم ديگر شبکه بيزين پويا ، شبکه بيزين دسته بندي چند متغيره در حلقه هاي بيزين مي باشد . مدلسازي شبکه بيزين پويا با وضعيت هاي مخفي به اندازه کافي اجازه يادگيري الگوهاي مختلف و هر ويژگي منحصر به فرد را در کلاسهاي ويديويي يا علايم سطح بالا نشان مي دهند . شکل زير يك معماري از شبکه هاي بيزين پويا را در سيستم هاي مبتني بر محتوا نمايش مي دهد [۹]

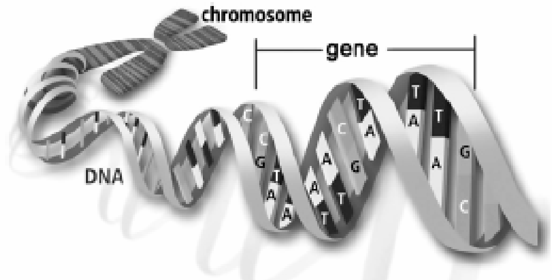
[۲۷] multijects
[۲۸] Dynamic Bayesian networks (DBNs)



معماري شبكه بيزين پويا براي سيستم بازيابي تصوير

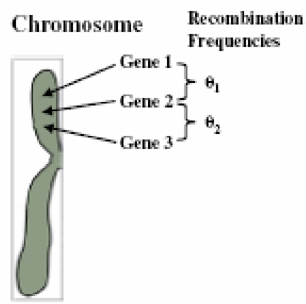
۳-۵- كاربرد شبكه هاي بيزين در تجزيه و تحليل اتصال (پيوستگي) ژنتيـك[۲۹] : تجزيـه و تحليـل اتـصال ژنتيك يك روش آماري براي نگاشت ژن ها بر روي كروموزومها مي باشـد ، و بـراي شناسـايي و پيـشگويي بيمـاري بسيار مفيد مي باشد . يكي از محدوديت هاي اين مسائل ، پيچيدگي محاسبات مي باشد . در اين مساله براي تجزيه و تحليل اتصال ژنتيك از شبكه هاي بيزين و براي آدرس دهي از روش هاي جديد استفاده شده است . مرتب سازي ژن ها بر روي كروموزوم و تعيين فاصله ما بين آنها در پيشگويي و تشخيص بيماري بسيار مفيد مي باشـد . قبـل از اين كه بيماري ظاهر شود (آشكار) گردد اجازه معالجه سريعتر را مي دهد و كشف بيماري از روي ژن ها مكان يابي شده و ژن هاي نزديك به آن بر روي كروموزوم مي تواند منجر به شناسايي افراد بـا احتمـال بـالا ي رخـداد شـود . بسياري از الگوريتم هايي كه براي تجزيه و تحليل اتصال ژنتيك به كار مـي رود الگـوريتم Eleston-Stewart يـا الگوريتم Lander-Green مي باشد . اولين الگوريتم در تعداد زيادي از افراد و تعداد معدودي از ژن ها و دومـين الگوريتم در تعداد كمي از افراد و تعداد زيادي از ژن ها انجام شد . بدون شـك تعـداد معـدودي از تجزيـه و تحليـل اتصال ژنتيك ايجاد شده اغلب از لحاظ سرعت پيوستگي ، طبقه بندي (تجزيه پـذير بـودن) و اتـصال بـزرگ قابـل ملاحظه مي باشند . تجزيه و تحليل اتصال ژنتيك ابزاري هستند براي حل مسائل در زمينه زيست شناسي كـه بايـد مد نظر قرار گيرند . سلولهاي انسان شامل ۲۳ جفت كروموزوم و ترتيب DNA تركيبي از ژن هاي منحصر به فرد و به ارث بردن والدين شان مي باشد . هر جفت شامل يك كروموزوم به ارث برده شده از پدر و يكي از مادر مي باشد . موقعيت هر يك از كروموزومها به جايگاه ژن دلالت مي كند . جايگاه ژن تابع ويژه اي كه به عنوان ژن شـناخته مـي شود . اين تابع مي تواند نتايج تركيب چندين ژن و همچنين تشخيص خون اشخاص را ارائه دهد . مراحـل واقعـي ژن ها نوع موروثي [۳۰] ناميده مي شود و نتايج قابل مشاهده انـواع مـوروثي (انـواع ژن) فنوتيـپ ناميـده مـي شـود . علامتگذار ژن ها موقعيت شناخته شده از دنباله DNA است كه مي تواند در جمعيت عمومي در هر شـخص يافـت شود . اين علامت ها ي به كار برده شده در موقعيت ژن هاي بيمار كمك مي كنند . شكل زير نمـايش كرومـوزوم ، تركيب DNA و تشخيص هويت ژن را نشان مي دهد . [۱۱]

۲۹ Genetic linkage analysis
۳۰ Genotype



دياگرام ژن بر روي كروموزوم

موقعيت كروموزرومها دنباله ايست ما بين دو والد كه به عنوان تقـاطع [۳۱] يـا وقـايع تركيبـي شـناخته شـده اسـت . فركانس تركيبي θ_i است كه تابع تركيبي ناميده مي شود و ما بين دو ژن متوالي به عنوان احتمال تركيبي وقايع مـا بين آنها تعريف مي شود . بنابراين اگر دو ژن غير متصل باشند يا غير مرتبط در اين صورت $\theta = 0.5$ مي باشد و ژن هاي متصل $\theta < 0.5$ مي باشد . شكل زير به شرح موقعيت سه ژن مي پردازد . [۱۱]

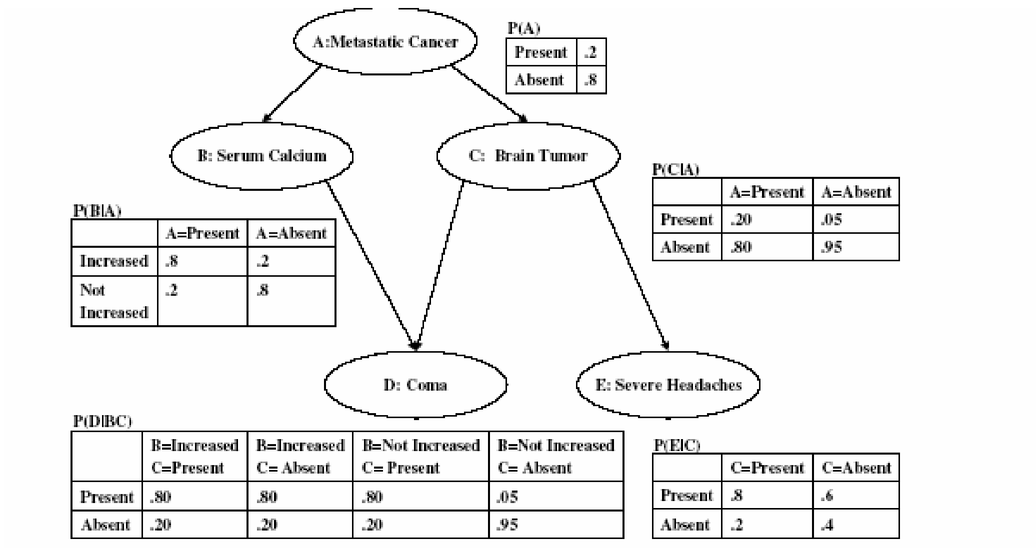


مرتب سازي ژن ها بر روي كروموزوم و تعريف فاصله مابين آنها

P براي نمايش جمعيت به صورت منحصر به فرد و e علامت ژنوتايپ (انواع موروثي) يا فنوتايپ مي باشد . $\hat{\theta}$ يك بردار شامل تركيب پي در پي ما بين هر جفت در ژن هاي متوالي است . از اين رو اگر ما n ژن داشته باشيم در ايـن صورت $\hat{\theta}$ شامل مقادير $n - 1\theta_i$ خواهد شـد . مـي تـوانيم $Pr(\mathrm{e}|P,\hat{\theta})$ را محاسـبه كـرده تـا احتمـال داده بـراي جمعيت و تركيب متوالي شناخته شود . وظيفه تجزيه و تحليـل اتـصال ژنتيـك محاسـبه احتمـال بـراي بردارهـاي چندگانه $\hat{\theta}$ مي باشد .

۳-۵-۱- شبكه هاي بيزين به كار برده شده : زماني كه حوزه تجزيه و تحليل اتـصال ژنتيـك بـا شـبكه بيـزين مدلسازي مي شود در اين صورت وظيفه آن محاسبه احتمال $Pr(e)$ براي نسبت هاي تركيبي مختلف است . شـبكه بيزين يك جفت (G, P) كه G گراف جهت دار غير حلقوي و P شامل مجموعه اي از فاكتورها مي باشد . هر نود G متناظر با متغير تصادفي براي مجموعه $\mathbf{X} = \{X_1, \ldots, X_n\}$ مي باشد و لبه ها به طور مستقيم براي نمايش مستقيم احتمال ما بين دو متغير متصل به كار مي رود . اين تاثير بوسيله پارامترهايي از شبكه P اندازه گيـري مـي شـود . براي هر $X_i \in \mathbf{X}$ والدين $\mathbf{Pa}_i$ وجود دارد كه P بايد شـامل فـاكتور $f(x_i, \mathbf{pa}_i) = Pr(x_i|\mathbf{pa}_i)$ باشـد . عمومـا ايـن فاكتورها در جدولي به نام جدول احتمالات شرطي مي باشد كه در اين صورت P شامل احتمالات هر يك از حـالات شرايط متغير بر روي نمونه هاي ممكن از والدين مي باشند . شبكه هاي بيـزين كامـل بـه صـورت توزيـع احتمـالي پيوسته بر روي متغيرهاي تصادفي تعريف مي شوند . تعريف گرافيكي يك شبكه بيزين ساده در شكل زير آمده است .

[۳۱] Crossover



شبكه بيزين

هر متغير تصادفي مجموعه اي از حالت هاي ممكن مي باشد براي مثال متغير C تومور مغز است كه مي تواند حالت "بود " يا "نبود " راغ نشان دهد و جدول احتمال شرطي انتساب شده به متغير با احتمالي نمايش داده شده است . و وضعيت والدين اتفاق افتاده متغير A است كه سرطان (غده) جابجا شونده مي باشد . در بسياري از شبكه هاي بيزين لبه ها نمي توانند احتمال موثر ما بين متغير ها را آشكار كنند اما آشكار شدن رابطه بين علت و معلول به صورت علت ◄ تاثير مي باشد . وظيفه شبكه بيزين ايجاد شده استنتاجو پاسخ به پرس و جو ها و محاسبه احتمالات مطلوب مي باشد . بسياري از الگوريتم هاي استنتاج مبتني بر كار اوليه مي باشند . شبكه بيزين ، محلي براي بسياري از استراتژي ها براي توليد طبقه هاي زدوده شده و ابزار مورد استفاده براي اين نواحي است كه مي تواند فوايد اين استراتژي ها را تحقيق كرده و به عنوان مزيت مدلسازي احتمالات به كار رود . [۱۱]

۳-۶- كشف دانش در پايگاه داده و داده كاوي و شبكه هاي بيزين : روش كشف دانش در پايگاه داده روش وسيع و توانمند در جستجوي دانش موجود در داده مي باشد . بنابراين هدف اصلي پيدا كردن الگوهاي معتبر از داده مي باشد . استخراج دانش از داده مي تواند به عنوان فرآيندي مد نظر قرار گيرد و مطابق با مراحلي از انتخاب و آماده سازي داده ، داده كاوي ، ارزيابي دانش استخراج شده مورد استفاده قرار گيرد . مرحله داده كاوي هسته اصلي در فرآيند كشف دانش از پايگاه داده مي باشد . اين مرحله به ايجاد مدلهاي مناسب براي نمايش الگوها و روابط مشخص در داده را شامل مي شود . نتايج اين مدلها توسط تحليل گر ارزيابي شده و با استفاده از مقادير پيش گويي شده براي ويژگيها با كاربر نهايي براي تعريف داده جديد استفاده مي شود . الگوريتم هوشمند محاسباتي براي داده كاوي مبتني بر شبكه هاي بيزين مي باشند . شبكه هاي بيزين تركيبي از چندين نود مي باشند كه هر نود نمايشي از شبكه يك متغير و يك ويژگي از پايگاه داده مي باشد كمانها متصل به هم بوده و جهت آنها دلالت به روابط وابسته به متغيرهايي است كه مي توانند آنها را تحت تاثير قرار دهند و جداول احتمال براي هر نود را شامل مي شوند . شبكه هاي بيزين مي توانند مدلهايي از روابط احتمالي ما بين متغيرها را تدوين كنند . اين مدلها شامل مولفه هاي كيفي براي نمايش وابسته ما بين نودها و مولفه هاي كمي براي جدول احتمالات شرطي باشند . بنابراين اين مولفه ها توزيع احتمالي پيوسته از X را در اين حوزه فراهم مي كنند . يكي از فوايد اصلي شبكه هاي بيزين در معناشناسي مي باشد كه نمايش سلسله مراتبي بين علت و معلول را آسان تر مي كند در نتيجه فرآيند تصميم گيري براي كاربر راحت تر خواهد شد . يكي از انواع شبكه هاي بيزين ، شبكه بيزين ساده مي باشد كه در اين شبكه همه صفات پايگاه داده به صورت رابطه دو طرفه مستقل مي باشند اما به نود پدر وابسته اند . بيزين ساده در مقابل

بسياري از روش هاي دسته بندي ساده و محاسبات موثرتري دارد . همچنين در مقابل بسياري از نويزهاي داده در مقابل ويژگيهاي نامرتبط مقاوم است و احتمالات ويژگيهاي ديگر را تحت تاثير قرار نمي دهد . [١٤]

٣-٧- بررسي اتصالات مغزي (مدلسازي سيستم عصبي) با استفاده از شبكه هاي بيزين پويا : براي مدلسازي اتصالات مغزي موثر از داده MRI مي توانيم از شبكه هاي بيزين پويا استفاده نماييم . زماني كه مدلسازي اتصالات مغزي صورت مي گيرد شبكه بيزين با نواحي مغزي فعال شده در ميان لبه هاي مشخص شده از فعل و انفعالات ما بين نواحي انتساب مي شوند . يك سيستم عصبي با به كار انداختن n ناحيه مغزي بايك سنسور با يك فرآيند يادگيري به صورت مجموعه $I=\{i : i=1, 2, \ldots n\}$ شاخص گذاري مي شود . فعاليت ناحيه مغزي با متوسط سري زماني fMRI اندازه گيري مي شود x_i فعاليت اندازه گيري شده در پاسخ به همودياليز ناخيه i مي باشد . توزيع احتمالي شبكه هاي بيزين بر روي نواحي مغزي فعال است . كه ساختار گرافيكي يك راه آسان از شرايط ويژه براي توزيع پارامترهاي موثر مي باشد . شبكه بيزين با ساختار S و توزيع پيوسته بر روي شرايط زماني $x=\{x_i : i\in I\}$ تعريف مي گردد . ساختار شبكه بيزين يك گراف جهت دار غير حلقوي است كه با فقدان سيكلهاي جهتي مشخص مي گردد . اگر a_i مجموعه فعال از والدين بر روي I باشد گراف جهت دار غير حلقوي يك روش ساده و منحصر به فرد براي تجزيه ضوابط احتمالي با احتمال شرطي زير را پيشنهاد مي كند :

$$P(x|\theta) = \prod_{i\in I} P(x_i|a_i, \theta_i)$$

$\theta=\{\theta_i : i\in I\}$ نمايش پارامترهاي احتمال شرطي مي باشد . شبكه هاي بيزين پويا مدل پيشرفته از شبكه هاي بيزين اند كه باعث بهم پيوستن مشخصات موقتي با سري زماني x مي گردد . به طور واضح نمايش فرآيندهاي موقتي از نواحي مغزي $x(t)=\{x_i(t): \quad i\in I\}$ براي نمايش تابع فعاليت n ناحيه مغزي در مدت زمان t مي باشد . براي مثال $t=1, 2, \ldots T$ متناظر با زمانهايي است كه از مغز اسكن شده و T تعداد كل اسكن ها را نمايش مي دهد . [١٨] به همين دليل مدل پوياي موقتي از فرآيند مغزي است كه ما مستلزم توزيع احتمالي بر روي مجموعه اي از متغير هاي تصادفي $\bigcup_{t=1}^{T} x(t)$ كه به طور تجربي پيچيده و گراف است ، مي شويم . با اجتناب از رشد پيچيدگي مدل فرض مي كنيم تغيرات موقت از تابع فعاليت مغزي بي حركت باشد و اولن ماركف به صورت $P(x(t+1)|x(t), \ldots x(1)) = P(x(t+1)|x(t))$ مي باشد . احتمال جابجايي به صورت مستقل بر روي t به صورت $P(x(t+1)|x(t))$ مي باشد . شبكه انتقال به صورت ساختار متصل ما بين دو اسكن مغزي پي در پي (متوالي) نمايش داده مي شود . توزيع پيوسته همه فرآيندهاي موقتي ملاقات مي شود. ساختار شبكه هاي بيزين پويا با حذف شبكه انتقال بر روي اسكن هاي متوالي براي همه $t=1, 2, \ldots T$ بدست مي آيد . شبكه بزين پويا قادر به مدلسازي شبكه برگشت پذير با قيود غير حلقوي در شبكه انتقال مي باشد . مهم ترين فايده مدلسازي سيستم شبكه با شبكه بيزين وجود سيكل تابعي در مغز به

صورت حلقه هاي متقارن پوستي [۳۲] است . اتصالات داخلي اسكن از نواحي مشابه مغزي به صـورت اتـصالات اوليـه مغزي مورد توجه قرار مي گيرد و اين پارامترها اجازه سازگاري را مي دهند . **[۱۸]**

۳-۸- موقعيت حركت ربات با استفاده از استنتاج شبكه هاي بيزين : در هـدايت ربـات هميـشه ربـات نمـي تواند موقعيت منحصر به فرد را تعريف كند . فقط با احساس اطلاعـات محلـي ، سنـسور آمـاده خطـاگيري و تغييـر كوچك ربات با نتايج دريافت شده است . بنابراين بسياري از روش هاي احتمالي پيـشنهاد شـده از عـدم قطعيـت و بهبود شديد موقعيت بر مي آيد . موقعيت حركت ربات شامل تخمين مناسب ربات در محيط مي باشد . بنـابراين دو كلاس موقعيت براي مساله وجود دارد موقعيت پيگردي [۳۳] و مكان يـابي سراسـري [۳۴] .در موقعيـت پيگـردي ربـات موقعيت اوليه را مي شناسد و فقط مستلزم كاهش ناپايداري در خواندن كيلومتر شمار [۳۵] مـي باشـد . اگـر موقعيـت اوليه شناخته نشود يا توسط ربات ربوده شود مساله يكي از موضع يابي هاي محلي را مانند حركت ربات بـه صـورت سراسري يا دنباله اي از فعاليت هاي دريافت شده را تخمين مي زند . دو ويژگي اصلي در سيستم وجـود دارد يكـي يكي استفاده از شبكه هاي بيزين براي نمايش شرايط وابسته روابط ما بين شـهود دريافـت شـده و فـصل مـشترك كريدور و مكان يابي سراسري اجرا مي شود . سيستم ساختار شبكه بيزين را از داده هاي محيطي كامل ياد مي گيرد . **[۲۳]**

۴- نتيجه گيري : شبكه هاي بيزين بيشتر در زمينه استدلالي به كار مي روند . اين شبكه ها روشي بـراي نمـايش توزيع احتمالي بزرگ به صورت نمايي و فشرده ميباشند كه اجازه محاسبات احتمالي بـه طـور مـوثر را مـي دهنـد . بديهي است شبكه هاي بيزين به استدلالهاي تحت شرايط نامشخص (احتمالي ، عدم قطعيت) كمك مي نمايند .

---

[۳۲] Cortico-Subcortical Loops
[۳۳] position tracking
[۳۴] Global Localization
[۳۵] odometer

matlab1.ir

۵- منابع :

[۱]- Masaharu Munetomo, Naoya Nurao, Kiyoshi Akama**. Introducing Assignment Ffunctions to Bayesian Optimization Algorithms.ELSEVIER**. ۶ **august ۲۰۰۷.**

[۲]-Donald e. Maurer, John p. Baker**. Fusing Multimodal Biometrics with Quality Estimates via Bayesian Belief Network.ELSEVIER ۲ august ۲۰۰۷.**

[۳]- M. Wiggins, A. Saad, B. Llitt, G. Vachtsevanos. **Evolving A Bayesian Classifier for Ecg-based age classification in medical applications.ELSEVIER.** received ۲۹ november ۲۰۰۵; **۲۰ march ۲۰۰۷.**

[۴]-Borja Calvo, Pedro Larran, Jose´ A. Lozan**o.LearningBayesian Classifiers From Positive and unlabeled examples . ELSEVIER ۱۵ August ۲۰۰۷**

[۵]- Nicandro Cruz-Ramireza, Héctor Gabriel Acosta-Mesa, Humberto Carrillo-Calvet, Luis Alonso Nava-FernJndez, Rocfo Erandi Barrientos-Martinez. **Diagnosis Of Breast Cancer Using Bayesian Networks: Acase Study .ELSEVIER**. **۱۵ february ۲۰۰۷.**

[۶]- Estevam R. Hruschka Jr, Nelson F.F. Ebecken**. Towards Efficient Variables Ordering for Bayesian Networks Classifier.ELSEVIER. ۲۰۰۷.**

[۷]- Olga Goubanova, Simon King**. Bayesian Networks For Phone Duration Prediction.. ۲۳ october ۲۰۰۷.**

[۸]- Patricio Garcia, Anal´a Amandi, Silvia Schiaffino, Marcelo Campo**. Evaluating Bayesian Networks Precision for Detecting Students Learning styles.ELSEVIER.** ۲۰۰۵; received in revised **november ۲۰۰۵.**

[۹]- Ankush Mittal, Krishnan v. Pagalthivarthi**, Temporal Bayesian Network Based Contextual Framework For Structured Information Mining.ELSEVIER. ۱۸ january ۲۰۰۷.**

[۱۰]- Kyung Jae Lee, Woojin Chang**. Bayesian Belief Network For Box-Office Performance: A Case Study on Korean Movies.ELSEVIER. expert systems with applications xxx (۲۰۰۷) xxx–xx.**

[۱۱]- David Allen, Adnan Darwiche**. Rc_Link: Genetic Linkage Analysis using bayesian networks.ELSEVIER. ۳ october ۲۰۰۷.**

[۱۲]- Mingyi Wang, Zuozhou Chen, Sylvie Cloutier**. A Hybrid Bayesian Network Learning Method for Constructing gene networks. ۱۳ august ۲۰۰۷.**

[۱۳]- Joe Frankel, Mirjam Wester, Ssimon King**. Articulatory Feature Recognition using dynamic bayesian networks.ELSEVIER.۱۸ march ۲۰۰۷.**

[۱۴]- Adamo l. de Santana, Carlos r. France^s, Cla´udio a. Rocha, Solon v. Carvalho, Nandamudi **l.** Vijaykumar, Liviane P. Rego, Joao C. Costa**. Strategies For Improving the Modeling and Interpretabilityt of Bayesian networks.ELSEVIER. ۱۳ november ۲۰۰۶.**

[۱۵]- Dan Wu**. Maximal Prime Subgraph Decomposition Of Bayesian Networks: a Relational Database perspective.ELSEVIER. ۱۶ November ۲۰۰۶ .**

[۱۶]-Andrzej Polanski, Joanna Polanska, Michal Jarzab, Malgorzata Wiench, Barbara Jarzab. **Application Of Bayesian Networks for inferring cause–effect relations from gene expression profiles of cancer versus normal cells. ۲۷ march ۲۰۰۷.**

[۱۷]- Seth Sullivant**. Algebraic Geometry of Gaussian Bayesian Networks.ELSEVIER. ۲۳ April ۲۰۰۷.**

[۱۸]- Jagath C. Rajapaksea,b,, Juan Zhoua**. Learning Effective Brain connectivity with dynamic bayesian networks.ELSEVIER**.۱٤ **june ۲۰۰۷.**

[۱۹]- Tshilidzi Marwala**, Bayesian Training of Neural Networks Using Genetic Programming.ELSEVIER.۲۷ March ۲۰۰۷.**

[۲۰]- Lili Sun, Prakash p. Shenoy**. Using Bayesian Networks for Bankruptcy Prediction: some Methodological Issues.ELSEVIER. ۱۲ June ۲۰۰٦.**

[۲۱]- Eitel j.m. Laur´a, Peter J. Duchessi**. A Methodology For Developing Bayesian Networks: An Application to Information Technology (it) Implementation.ELSEVIER. ۱۰ March ۲۰۰٦.**

[۲۲]- Reyes Pavo´ N, Fernando D´ Azb, Victoria Luzo´ n**. A Model for Parameter Setting Based On Bayesian Networks.ELSEVIER. ۲ february ۲۰۰۷.**

[۲۳]- Hongjun Zhou, Shigeyuki Sakane**. Mobile Robot Localization Using Active Sensing Based On bayesian Network Inference.۱۹ December ۲۰۰٦.**

[۲٤]- Sinisa ToDorovic, Michael C. Nechyba**. Interpretation Of Complex Scenes Using Dynamic Tree-Structure Bayesian Networks.ELSEVIER. ۲۱ december ۲۰۰٦.**

[۲٥]- Craig s. Galbraith, Alex f. Denoble, Sanford b. Ehrlich, Doug m. Kline**. Can Experts Really Assess Future Technology Succes?Neural Network and Bayesian Analysis of Early Stage Technology Proposals.ELSEVIER.۲۰۰۷.**

[۲٦]- Ana c.v. De Melo a,*, Adilson J. Sanchez. **Software Maintenance Project Delays Prediction Using Bayesian Networks.ELSEVIER.۲۰۰۷**.

[۲۷]- Eitel J.M. LaurA A,., Peter J. Duchessi. **A Bayesian Belief Network for it Implementation decision support.ELSEVIER. ۹ march ۲۰۰٦.**

[۲۸]- Abdelaziz Oualia,B,., Amar Ramdane Cherifb, Married -Odile Krebsa**. Data Mining Based Bayesian Networks For Best Class. Elsevier. ۱۲ October ۲۰۰٥.**

[۲۹]- B. Naticchia A, A. Fernandez-Gonzalez B, A. Carbonari**. Bayesian Network Model For the Design Of Roofpond Equipped Buildings. ELSEVIER. ۱۱ july ۲۰۰٦.**

[۳۰]- Martin Neila,b, , Manesh Tailorb, David Marqueza, Norman Fentona,B, Peter Heart. **Modelling Dependable Systems Using Hybrid Bayesian Networks.ELSEVIER.۲۰۰۷.**