سیزدهمین کنفرانس ملی انجمن کامپیوتر ایران
جزیره کیش، خلیج فارس، ایران ۱۹ الی ۲۱ اسفند ۱۳۸۶
پوستر

استفاده از یادگیری تقویتی در مساله کاهش ابعاد.

مریم باستانفرد

دانشگاه آزاد اسلامی سپیدان

bastanfard@yahoo.com

چکیده در این مقاله، شیوه ای جهت انتخاب فضای ویژگیها و کاهش ابعاد، با استفاده از یادگیری تقویتی ارائه می‌شود. به این منظور از یک سیستم چندعامله استفاده شده که در ان عاملهای یادگیرنده می کوشند تا با مواجهه با مساله و همکاری با یکدیگر به یک هدف مشترک، که به دست آوردن مجموعه ای ازبهترین ویژگیها میباشد، برسند. یادگیری عاملها به واسطه بررسی میزان تاثیر گروهی از ویژگیها ، براساس بیشترین جداپذیریی که در بین داده ها ایجاد می کنند، با سعی و خطا خواهد بود.این جدا پذیری مبتنی بر روش جداپذیری فیشر و استفاده از ضریب همبستگی در جهت بیشینه کردن حداقل فاصله کلاسها می باشد.

کلید واژه- فضای ویژگی، عامل، یادگیری تقویتی، جداپذیری فیشر، ضریب همبستگی.

1- مقدمه

در بسیاری از موارد، وابستگی مهم ویژگیهای مختلف و چگونگی تاثیر متقابل آنها را نمی دانیم. قضاوت درباره کارایی اطلاعاتی که توسط یک ویژگی بیان میشود، بسیار سخت است. ممکن است به مسایلی برخورد کنیم[1]، که هرگز نمی توان به تعداد کافی داده بدست آورد تا با استفاده از آنها بتوان سیستمی را مدل کرد که تمام موارد نادر را در بر بگیرد؛ چون داده ها نادر هستند.

بنابراین اولین وظیفه هر سیستم کلاسه بند، آنست که تصمیم بگیرد کدام یک از تعداد بسیار زیاد ویژگی های در دسترس، برای این منظور خاص استفاده شوند. پیشنهاد ویژگیها کار راحتیست، اما اینکه هر کدام تا چه اندازه سودمند هستند، بحث دیگری است. برای این منظور از ویژگیهایی استفاده می شودکه توانایی لازم در جدا کردن کلاسهای مختلف را از هم داشته باشند، این در حالیست که تعداد زیاد از حد این ویژگیها باعث افزایش ابعاد و پیچیدگی مساله خواهد شد. به عبارتی ویژگیهایی که اطلاعات جدیدی به دست نمی دهند یعنی اطلاعاتی مشابه به اطلاعات حاصل از سایر ویژگیها را در خود دارند، ارزش زیادی ندارند. بنابراین به مجموعه ویژگیهایی نیاز است که تا حد امکان جامع و مانع باشند. در مجموع، منظور از انتخاب ویژگیها گزینش گروهی از ویژگیهاست به گونه ای که با حفظ شباهت اعضا در یک کلاس خاص، در حد امکان کلاسهای مختلف از هم قابل تفکیک باشند. این ابعاد باید به گونه ای باشد که اولا، اطلاعات از دست رفته در روند کاهش، باید در حداقل ممکن باشد. ثانیا، فضای حاصل شده، باید توانایی لازم در جهت جدا سازی را داشته باشد[3].

در اینجا با تکیه بر یادگیری تقویتی[12] الگوریتمی پیشنهاد میشود که در آن عاملهایی به صورت غیر متمرکز[10]، بر اساس معیاری مبتنی بر معیار فیشر و ضریب همبستگی در طی یک فرایند هیوریستیک هوشمند به انتخاب ویژگیها و ابعاد میپردازند.

2- معیارهای جدا پذیری

یک تدبیر آنست که ویژگیهایی انتخاب شوند که همبستگی نیرومندی با یکدیگر داشته باشند، که با عنوان

Archive of SID

www.SID.ir

بیشترین ارتباط معروف است. از دیدگاه دیگر میتوان ویژگیهایی را انتخاب کرد که در تقابل با یکدیگر باشند در عین حال که بیشترین ارتباط را با هم دارند، این تدبیر، با نام حداقل زوائد_حداکثر ارتباط شناخته شده است. راهکار اخیر بسیار از روش بیشترین ارتباط کارا تر است[7]. معیار اطلاعات متقابل [1] میتواند در جهت بیان اثر وابستگی بکار رود. به عنوان یک نمونه خاص، همبستگی و ارتباط و حداقل زواید[4] میتوانند به خوبی با ضریب همبستگی و معیار فیشر لحاظ شوند.

در اینجا نشان داده میشود که برای به دست آوردن چنین مجموعه ای می توان از یک سیستم با چند عامل یادگیرنده استفاده کرد که در این سیستم یادگیری عاملها بر پایه همکاری تا رسیدن به بهترین ترکیب از ویژگیها، مبتنی بر یادگیری تقویتی انجام میشود. در این روش با پیشنهاد ترکیبهای مختلف ویژگیها توسط عامل و دریافت جایزه بر اساس عملکرد مجموعه ویژگیها، بهترین مجموعه از ویژگیها انتخاب میشود. بهترین مجموعه بسته به مورد، می تواند متفاوت باشد. این تفاوت ناشی از گونه های مختلف آرایش نمونه های آموزشی است. این آرایش برتری و مزیت استفاده از گروهی از ویژگیها را نسبت به بکارگیری عده دیگری را باعث می شود. برتری عده ای به دیگران بر اساس میزان جدا پذیری مناسبی است که در بین کلاسهای مختلف ایجاد می کنند. به علاوه، یک ویژگی می تواند به خودی خود ناکار آمد بوده و جدا سازی مناسبی ایجاد نکند، اما این امکان وجود دارد که همراه با دیگر ویژگیها کارایی بالایی داشته باشد و جدا سازی قدرتمندی ایجاد کند. بنابراین، ویژگیها به تنهایی قدرت جدا سازی زیادی ندارند. استفاده همزمان آنها، جدا سازی مناسبی را سبب می شود.

2-1- معیار جداپذیری فیشر

معیار فیشر روشی است که میتواند اطلاعاتی در مورد کلاسها زمانی که با مجموعه ای از ویژگیها نشان داده میشوند را بیان می کند .در این روش فرض میشود که هر یک از کلاسها که هر کدام مجموعه ای از نقاط را شامل میشود، بوسیله میانگین و واریانسش بیان میشود (برای ابعاد بیشتر از 1 از ماتریس کوواریانس استفاده می شود). کوچکترین واریانس در داخل کلاسها و بزرگترین فاصله در بین آنها ، پیامد بهترین ترکیب از ویژگیها خواهد بود.

بنابراین بهترین مجموعه از ویژگیها دارای بالاترین میزان برای نسبت زیر میباشد.

$$\frac{\text{inter-class variance}}{\text{intra-class variance}}$$

از آنجا که واریانس در داخل کلاسها میتواند متفاوت باشد، میتوان این مورد را جایگزین کرد

$$\frac{\text{variance between classes}}{\text{higher of the intra-class variance}}.$$

واریانس داخلی کلاسها به صورت کوواریانس میانگینها محاسبه نمی شود بلکه این مقدار کوواریانس همه نقاط میباشد[15].

به طور کلی، اگر داده هایی در n کلاس فرض شوند و هر کلاس با ماتریس X_i نشان داده شده که در آن، ستونها بردارهای ویژگی هستند، آنگاه داریم:

$$A = \mathrm{cov}([X_1 X_2 X_3 ... X_n])$$
$$B_i = \mathrm{cov}(X_i)$$
$$\lambda_i = \max(eig(B_i^{-1} A)) \quad (1)$$

بنابراین معیار فیشر با ریشه دوم کوچکترین λ_i برابر است. هر چه این مقدار بزرگتر باشد نشان دهنده جداپذیری بیشتری است. رابطه (1) در [15] محاسبه شده است.

2-2- ضریب همبستگی

یکی از شاخصهای مهم آماری ضریب همبستگی است که به منظور بررسی میزان وابستگی و استقلال داده ها بکار میرود. این معیار با این معیار با استفاده از رابطه (2)، برای داده های یک بعدی محاسبه میشود.

$$\rho(i) = \mathrm{cov}(x_i, y) / \sqrt{\mathrm{var}(x_i)\,\mathrm{var}(y)} \quad (2)$$

که در آن $\mathrm{Cov}(x,y)$ کوواریانس x و y و δ_x واریانس x است. ضریب همبستگی در مسایل با بیشتر از یک بعد با استفاده از ماتریس کوواریانس، برای تک تک ابعاد محاسبه میشود.

از این معیار میتوان به منظور مقایسه ویژگیها استفاده کرد. ضریب همبستگی، وابستگی خطی را با عددی در محدوده [1و0] بیان میکند. مقدار 1 نشان دهنده آنست که تابع غیر صفری وجود دارد که بی هیچ ابهامی یک نمونه را به کلاسی نسبت میدهد و در مقابل مقدار 0 عدم وابستگی را نشان میدهد.

Archive of SID

www.SID.ir

3- سیستم یادگیری

3-1- عامل

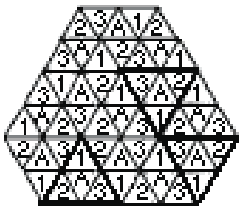
در اینجا هر عامل درک شخصی خود را از پیرامون خود دارد و چون دید عاملها محدود است در نتیجه در طول مسیر، برداشت شخصی خود را بدون در نظر گرفتن سایرین از محیط اطرافش دارد.

عامل یک حالت داخلی جاری دارد و اعمال را برای تغییر حالت جاری می تواند اجرا کند. در هر گام حالت جاری ذخیره می شود، عمل اجرا می شود و حالت بعدی ذخیره می شود. مدل محیطی حالت جاری را ذخیره می کند و انتقال های حالت را پیاده سازی می کند. خود عامل از مسأله یادگیری مجزاست و مجموعه ای ازعمل هایی که می تواند انتخاب کند، دارد و از سیاست آمده از یک شئ کنترل کننده پیروی می کند.

3-2- الگوریتم یادگیری

مجموعه تمام ویژگیهای مورد نظرکه از داده ها حاصل میشود، به صورت مجموعه های 3 تایی تقسیم میشود. این تقسیم بندی میتواند کاملا تصادفی باشد. تقسیم به مجموعه های سه تایی باعث میشود که تعداد زیرمحموعه ها کاهش یابد. به علاوه اینکه پس از چندین اپیسود با تکیه بر استنباط از تعداد انتخابهاکاسته شود، بنابراین حالتهای کمتری برای بررسی وجود خواهد داشت. در صورتی که با تقسیم به مجموعه هایی با تعداد اعضای بیشتر، از آنجایی که حالتهای انتخاب عمل، بسیار بیشتر میشوند به تعداد اپیسودهای خیلی بیشتری برای یادگیری لازم است که زمان بیشتری را صرف میکند.

به ازای هر مجموعه یک عامل در نظر گرفته میشود که هر کدام، تعدادی از این ویژگیها را پیشنهاد میدهد و بر اساس دریافت جایزه یاد می گیرد. در هر نوبت یک عامل میتواند یک ویژگی را انتخاب کند و یا بی هیچ انتخابی به حالت بعدی برود (شکل شماره 1).



شکل شماره 1: ساختار محیط و عملهای انتخابی برای هر عامل.

این توصیف یک ماشین غیر قطعی_λ است که بدون انجام هیچ عملی نیز میتواند تغییر حالت دهد [9]. با اضافه کردن سه عضو به صورت عضو اضافی به هر یک از زیر مجموعه ها، می توان آنرا به یک ماشین متناهی تبدیل کرد.

$T_i = F_i \cup \{D_1, D_2, D_3\}$

که در اینجا F_i زیر مجموعه ای 3 تایی از ویژگیهاست و D_i عضو اضافی شماره i اضافه شده به این زیرمجموعه است.

درپروسه یادگیری هر عامل، که درکی اختصاصی از محیط دارد، اپیسودها که از یک نقطه تصادفی شروع و تا پایان، که حالت جذب است ادامه می یابد، به عنوان نمونه آموزشی بکار میرود. هرعامل به صورت غیر متمرکز، یکی از اعضای مجموعه خود را بر اساس روش soft max [12][13]، انتخاب و به حافظه اپیسودیکش می افزاید، در انتها مجموعه پیشنهادی همه عاملها با هم بررسی شده و بر اساس معیار فیشر (مبتنی بر بردار و مقادیر ویژه ماتریس کوواریانس)، ضریب همبستگی، میزان جداپذیری داده ها بر اساس این مجموعه ویژگیها سنجیده شده و به انتخابگرهای آنها پاداش داده میشود. در این بررسی همه انتخابها در یک سطح قرار دارند به عبارت دیگر ترتیب انتخاب یک عمل در میزان جایزه ای که به آن تعلق می گیرد، اهمیتی ندارد بلکه صرفا وجود یا عدم وجود یک ویژگی در مجموعه نهایی مورد بررسی قرار میگیرد. بنابراین بر خلاف روش مرسوم در یادگیری Q، که جایزه به صورت کاهشی و تنها به انتخاب در نقطه آغازی اپیسود تعلق میگیرد[2][14]، در اینجا بر اساس رابطه (3)، جایزه به تمام انتخابهای موجود در یک اپیسود، صرفا بر پایه تاثیرشان اختصاص می یابد. در اینجا همه زوجهای حالت_عمل در یک اپیسود، در یک سطح قرار دارند و تنها تاثیر آنها در میزان جداپذیری در بین نمومه

Archive of SID

www.SID.ir

هاست که به هر یک از آنها ارزش خاصی میدهد، نه ترتیب انتخاب آنها.

$$R_i = \lambda^N_{\max} \times \cos(\alpha_i) \times \gamma^n \times \rho_i \quad (3)$$
$$\lambda^N_{\max} = \lambda_{\max} - \bar{\lambda} / \lambda_{\max} - \lambda_{\min}$$
$$\rho_i = \max_j \rho_{ij}$$

در این روابط R_i پاداش مربوط به هر یک از ویژگیهای انتخابی در این اپیسود، بر اساس مقدار تاثیرش است و λ^N_{max} بزرگترین مقدار ویژه ماتریس کوواریانس، α_i زاویه بزرگترین بردار ویژه با ویژگی i ، γ عددی بین صفر ویک، n تعداد کل ویژگیهای انتخاب شده در این اپیسود، ρ_{ij} ضریب همبستگی داده ها بواسطه ویژگیهای i ام و j ام انتخابی در این اپیسود میباشد.

در حالت نهایی ، برای عامل(ها)یی که همه انتخابهایش در این اپیسود ، عضو اضافی بوده است، پاداش3+ و پاداش R_i برای ویژگیهای انتخابی، برای هر عامل منظور میشود.

پس از تکرار روند فوق، در نهایت در هر یک از زیرمجموعه های 6 عضوی T_i، انتخابها ارزششان به میزان واقعی همگرا شده [6] و انتخاب 3 عضوی که دارای بیشترین ارزش هستند، از هر زیرمجموعه، بهترین مجموعه ویژگیها را نتیجه میدهد.

4- نتایج آزمایشی

برای بررسی نتیجه و کارایی الگوریتم پیشنهادی انتخاب مجموعه ویژگیها، از یک مجموعه داده های واقعی مربوط به تصاویر سونوگرافی بافت تیرویید استفاده شده است.

سونوگرامهای مربوط به بافت غده تیرویید با دو نوع بافت متفاوتء که در مقایسه چشمی به صورت صد در صد قابل تشخیص نیستند، توسط یک سیستم کلاسه بندی، با تکیه بر مجموعه ویژگیهای مناسب، قابلیت جداپذیری دارند. این بافتها عباتند از بافت نرمال و بافت متورم که به افتخار پزشک ژاپنی کاشف این آسیب، به نام ایشان هاشیموتو نام گرفته است. بیماری هاشیموتو یا لیمفوسیتیک (LT)، که در مورد خانمها پنج برابر بیشتر از آقایان دیده میشود، زمانی مطرح است که تیرویید نمی تواند در زمان مشخص به میزان کافی هورمون ترشح کند. نمونه ای از دو بافت نرمال و آسیب دیده،که از تصاویر سونوگرافی به صورت دستی جدا شده اند، در شکل شماره 2 دیده میشود.

در اینجا سعی بر اینست مجموعه ویژگیهایی انتخاب شود که حداکثر میزان جدا پذیری بافت تصویر قسمتی از سونوگرام را، که دستی جدا شده است، ممکن گرداند. تعداد نمونه های آموزشی از هر دو نوع داده در جدول 1 نشان داده شده است.

شکل شماره 2: نمونه هایی از بافت تیروئید. الف | ب

الف: نمونه ای از بافت نرمال، ب: نمونه ای از بافت آسیب دیده.

تصاویر مورد استفاده به عنوان نمونه های آموزشی، با دستگاه سونوگرافی از نوع Toshiba ECCOCEE (console model SSA-340A, transducer model PLF-805ST at frequency 8MHz). تصویر برداری شده اند. این تصاویر با 8 بیت قدرت تفکیک و 256 سطح خاکستری ذخیره شده اند.

جدول 1: تعداد داده ها از هر نوع.

نوع بافت تیروییید	تعداد نمونه ها
بافت نرمال	165
بافت آسیب دیده	175

الگوریتم با مجموعه اولیه ویژگیها شامل 42 عضو، برای رسیدن به مجموعه بهینه، جامع و مانع[5][8]، ویژگیها اجرا شده است. این ویژگیها عبارتند از:

میانگین، واریانس، آنتروپی، دهک اول هیستوگرام، ده ضریب اول تبدیل فوریه، مقدار انرژی در هر یک از زیر باندهای تبدیل موجک در 3 سطح و 9 مقدار ویژگی هارالیک (جدول 2). برای محاسبه ویژگیهای هارالیک از ماتریس کواکیورنس، به ازای بردار فاصله ای معادل d=[11,0] استفاده شده است. در این بردار فاصله، عرض از مبدا مساوی با صفر، با ذات تصویر برداری سونوگرافی، که در آن تابش به صورت عمودی انجام می شود، کاملا هماهنگی دارد.

الگوریتم با $\gamma = 0.5$ برای داده ها و مجموعه اولیه ویژگیها اجرا شده است. عاملها 500 اپیسود را به عنوان نمونه آموزشی دیده و یاد گرفته اند، در نهایت عاملها پس از پیمودن روند یادگیری، به مجموعه ای از ویژگیها به عنوان

Archive of SID

www.SID.ir

جواب میرسند، که تا حد امکان توانایی جدا کردن این داده ها را دارد. این مجموعه جواب عبارتست از مقدار انرژی در هر یک از زیر باندهای تبدیل موجک در 3 سطح(10 مقدار)، انتروپی بافت و ارتباط بافت از ویژگیهای هارالیک.

مجموعه ویژگیهایی که برای همین داده ها بر اساس روش پیشرو انتخاب ویژگی مبتنی بر معیاری مشابه، برای رسیدن به دقتی یکسان با روش متکی بر یادگیری تقویتی، حاصل شد عبارتست ازسه مقدار اول هیستوگرام ، انتروپی بافت و ارتباط بافت و یکنواختی انرژی بافت از ویژگیهای هارالیک، مقدار انرژی در هر یک از زیر باندهای تبدیل موجک در 3 سطح(10 مقدار).

جدول2: ویژگیهای هارالیک.

	نام	توضیح
H1	texture cluster tendency	$\sum_{ij}(i-\mu_i+j-\mu_j)^2 C_{ij}$
H2	texture entropy	$-\sum_{ij} C_{ij}\log C_{ij}$
H3	texture contrast	$\sum_{ij}\lvert i-j\rvert C_{ij}$
H4	texture correlation	$\frac{\sum_{ij}(i-\mu_i)(j-\mu_j)C_{ij}}{\sqrt{\mathrm{var}(i)\mathrm{var}(j)}}$
H5	texture homogeneity	$\sum_{ij}\frac{C_{ij}}{1+\lvert i-j\rvert}$
H6	texture inverse difference moment	$\sum_{ij,i\neq j}\frac{C_{ij}}{\lvert i-j\rvert}$
H7	maximum texture probability	$\max_{ij} C_{ij}$
H8	texture probability of run length of 2	$\sum_{i}\frac{(C_i-C_{ii})^2 C_{ii}}{C_i^2}$
H9	uniformity of texture energy	$\sum_{ij} C_{ij}^2$

همانطور که دیده میشود با آنکه با تعداد کمتر ویژگی نیز میتوان به همین میزان جداپذیری رسید در روش دوم گروهی ویژگی با اطلاعات غیر مستقل و قابل بیان با سایر ویژگیها نیز به مجموعه جواب تحمیل شده اند.

این بدان دلیل است که در روشهایی مانند پیش رونده و پس رونده که در آنها شانس ویژگیها برای انتخاب در هر مرحله کاملا وابسته به ویژگی انتخاب شده در مرحله قبل است. به علت وابستگی روش به ترتیب انتخاب ویژگیها و عدم امکان واپس نگری در هر مرحله، که هیچگاه عدم وجود(وجود) یک ویژگی که قبلا اضافه (حذف) شده، در سطوح بالاتر، دوباره بررسی نمیشود؛ که میتواند باعث همپوشی اطلاعات حاصل از ویژگیها باشد. همچنین به دلیل نگاه و بررسی مقطعی به بهترین ویژگی به صورت محلی، بدون در نظر گرفتن نتیجه مجموع، در نهایت مجموعه جواب یک بهترین محلی خواهد بود.

معیار فیشر داده ها را دارای تابع توزیع گوسی فرض میکند که میتواند با مقادیر میانگین و واریانس، نمایش داده شود. این معیار برای مجموعه ویژگیهای حاصل از بکارگیری الگوریتم پیشنهادی 16.63 و برای مجموعه ویژگیهای منتج از الگوریتم پیشرونده برابر با مقدار 15.22 است. ضریب همبستگی نیز که میزان وابستگی یک نمونه جدید را به کلاسهای مختلف را بررسی میکند برای هر یک از مجموعه های ویژگی به ترتیب مقادیر0.71 و 0.56 است. هر دو این معیارها نشان دهنده برتری مجموعه ویژگی جدید است. در جدول شماره 3 نتایج حاصل از دو سیستم که بر اساس ویژگیهای به دست آمده از دو روش مختلف انتخاب ویژگیها، الگوریتم پیشرونده و الگوریتم پیشنهادی متکی بر یادگیری تقویتی، در کلاسه بندی داده ها با استفاده از الگوریتم K نزدیکترین همسایه، با هم مقایسه شده اند.

جدول 3: مقایسه نتایج کلاسه بندی مبتنی بر الگوریتم پیشنهادی و الگوریتم پیشرونده.

	الگوریتم پیشنهادی	الگوریتم پیشرونده
بافت آسیب دیده	٪95.49	٪87.3
بافت نرمال	٪96.85	٪92.6

از میان کل داده های موجود، در هر نوبت، تعدادی به عنوان نمونه آموزشی و بقیه به عنوان نمونه آزمایشی به صورت تصادفی انتخاب شده و روند دسته بندی برای آنها، بر اساس دسته بندی K نزدیکترین همسایه، انجام میشود. این روال چندین مرتبه تکرار و میانگین محاسبه میشود. بر اساس نتایج بدست آمده، تشخیص بافتهای نرمال و بافتهای آسیب دیده با تکیه بر مجموعه ویژگی حاصل از الگوریتم پیشنهادی، به ترتیب برابر با 96.85٪ و 95.49٪ است.

5- نتیجه گیری

در این مقاله به بررسی یک سیستم چند عامله پرداخته شد

Archive of SID

www.SID.ir

که در آن عاملها بر اساس یادگیری تقویتی سعی در هماهنگ کردن رفتارشان با یکدیگر در جهت دستیابی به یک هدف کلی و واحد به منظور یافتن مجموعه ای از ویژگیها برای بیان و نمایش داده ها با حداقل بعد به صورتی که حداقل اطلاعات از دست برود، دارند. به این منظور الگوریتم جدیدی، مبتنی بر یادگیری تقویتی با تکیه بر معیارهای مناسب، پیشنهاد شد که با توجه به هوشمندی در به کار بردن هیوریستیک بررسی احتمالات مختلف بر خلاف روشهای موجود، در زمان اجرای مناسب بسیار کاراتر عمل میکند.

مراجع

[1] A. Berger, S. Della Pietra, V. Della Pietra. A maximum entropy approach to natural language processing. In Computational Linguistics, 22, 1996, pp1-36.
[2] D. Dutech, Solving POMDP Using Selected Past-Events, Proc the 14th European Conference on Artificial Intelligence, ECAI2000, 2000.
[3] H.C. Peng, F. Long and Ding, C., "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 8, pp.1226-1238, 2005.
[4] J. Reunanen. Overfitting in making comparisons between variable selection methods. JMLR, 3:1371–1382 (this issue), 2003.
[5] J. Smith and S.-F. Chang, “Transform features for texture classification and discrimination in large image database”, IEEE Intl. Conf. on Image Proc., 1994.
[6] K. Miyazaki; M. Yamamura; and S. Kobayashi, On the rationality of profit sharing in reinforcement learning. In proceedings of the 3rd international conference on fuzzy logic, Neural nets and soft computing, 1994.
[7] K. Torkkola. Feature extraction by non-parametric mutual information maximization. JMLR, 3:1415–1438, 2003.
[8] M. Stricker and M. Orengo, "Similarity of Color Images", Proc. SPIE Storage and Retrieval for Image and Video Databases, 1995.
[9] M. Littman, A. Cassandra , L.Kaelbling, Learning Policies for Partially Observable Environments: Scaling Up, Proc. the 12th Int. Conference on Machine Learning, 1995.
[10] O. Buffet, A. Dutech , F.Charpillet, Multiagent Systems by Incremental Reinforcement Learning, LORIA, BP, 2002.
[11] P.L. Lanzi, S.W. Wilson, "Optimal classifier system performance in non-markov environments", IlliGAL Technical Report 99022, University of Illinois, 1999.
[12] R.S. Sutton,A.G. Barto, Reinforcement learning, Bradford Book, MIT Press, Cambridge, MA, 1998.
[13] S.B. Thrun, "The role of exploration in learning control with neural networks", In D. A. White and D. A. Sofge, editors, Handbook of Intelligent Control: Neural, Fuzzy and Adaptive Approaches, Florence, Kentucky, Van Nostrand Reinhold, 1992.
[14] S. Singh, T. Jaakkola , M. Jordan , Learning Without State Estimation in Partially Observable Markovian Decision Processes, Proc. The Eleventh International Conference on Machine Learning, 1994.
[15] Z. Kotek, I. Br°uha, V. Chalupa, and J. Jel´ınek. Adaptivn´ı a u˘c´ıc´ı se syst´emy. SNTL, 1980.

Archive of SID

www.SID.ir