

عنوان مقاله:

پیش بینی پروتئینهای متصل شونده به لیپید بر اساس ماشین بردار پشتیبان با استفاده از خصوصیات مربوط به توالبها

محل انتشار:

چهارمین همایش بیوانفورماتیک ایران (سال: 1391)

تعداد صفحات اصل مقاله: 2

نویسندگان:

محمد رضا بختیار زاده - گروه علوم دامی، دانشگاه تهران، کرج، ایران؛ گروه زراعت و اصلاح نباتات، دانشگاه شیراز، شیراز، ایران

محمد مرادی شهر

بابک اسماعیل ابراهیمی

خلاصه مقاله:

پیش بینی عملکرد پروتئینهای کاندید و جدید یکی از وظایف مهم و پیچیده بعد از دوره پس-ژنومی میباشد. با این حال شناسائی پروتئینهای با شباهت کم و یا بدون شباهت با پروتئینها یا پپتیدهای با عملکرد شناخته شده یکی از مشکلات بسیار مهم میباشد. روشهای یادگیری ماشین، مانند بردار ماشین پشتیبان (SVM) از روشهای مفید برای پیش بینی چنین پروتئینهایی با توالبهای پروتئینی متنوع میباشد. پروتئینهای متصل شونده به لیپید (LBPs) نقشهایی حیاتی در سلولهای مختلف همانند علامت دهی، عبور و مرور غشائی، تنظیم، پاسخ ایمنی، متابولیسم لیپید و انتقال ایفا میکنند. با توجه به وجودتنوع در توالبهای پروتئینی بین LBPs، برای پیش بینی این پروتئینهای نیاز به روشهایی احساس میشود که بر پایه شباهت بین توالبهای نباشند. در این مطالعه از یک روش بر پایه SVM و استفاده از خصوصیات فیزیکی و شیمیائی (مانند خصوصیات خودهمبستگی، ترکیب، انتقال و توزیع، درجه-توالی-کواسی و اسید آمینه کاذب) و ترکیب اسید آمینه های توالبهای پروتئینی به منظور پیش بینی و طبقه بندی LBPs استفاده گردید. مجموعه داده های مورد استفاده در این مطالعه برای آموزش و امتحان مدل از سایت UniProt استخراج گردید. نرم افزار LibSVM (یک نرم افزار برای پیاده سازی SVM) به منظور طبقه بندی پروتئینهای LBPs استفاده گردید. مجموعه دادههای LBPs شامل 10603 توالی پروتئینی بود که متعلق به 9 گروه شامل تجزیه لیپید، متابولیسم لیپید، تولید لیپید، انتقال لیپید، اتصال به لیپید، تولید لیپوپلیساکارید، لیپوپروتئین، لیپویل و همه پروتئینهای متصل شونده به لیپید میباشد و نقش مهمی در کنترل اعمال سلولی ایفا میکنند. مجموعه داده مربوط به پروتئینهای غیر مرتبط با LBPs (non-LBPs) حاوی 185628 توالی پروتئینی بود. در این تحقیق از روش طبقه بندی دوتائی برای همه گروهها استفاده گردید. هر کدام از گروههای LBPs شامل یک زیرمجموعه داده مثبت (گروه های LBPs) و یک زیرمجموعه داده منفی (non-LBPs) بود. شباهت بین توالبهای پروتئینی در همه مجموعه داده ها کمتر از 90% تعیین گردید. پنج زیرمجموعه داده منفی به طور تصادفی از مجموعه دادههای non-LBPs برای بررسی هر گروه به منظور اجتناب از اربیب ناشی از انتخاب زیر مجموعه داده های منفی در نظر گرفته شد. نتایج پیش بینی ها با استفاده از آزمونهای تائید-مقاطع پنج تائی و مجموعه داده مستقل مورد ارزیابی قرار گرفت. این روش LBPs و non-LBPs را با صحت 89.28% بر اساس آزمون تائید-مقاطع پنج تائی طبقه بندی کرد. صحت پیش بینی محاسبه شده بر پایه آزمون تائید-مقاطع پنج تائی برای گروههای شامل تجزیه لیپید، متابولیسم لیپید، تولید لیپید، انتقال لیپید، اتصال به لیپید، تولید لیپوپلیساکارید، لیپوپروتئین و لیپویل به ترتیب برابر با 92.25، 89.15، 94.74، 88.84، 90.61، 93.26، 89.06 و 98.58% برآورد شد. با استفاده از آزمون مجموعه داده مستقل LBPs با صحت 89.55% نسبت به non-LBPs شناسائی گردیدند. صحتهای برابر با 90.85، 09، 95.36، 91.06، 3.923، 93.15، 90.10 و 99.55% به ترتیب برای گروههای مذکور و بر اساس آزمون ...

کلمات کلیدی:

پروتئینهای متصل شونده به لیپید، SVM، پیش بینی

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/287581>



