ملخص الرسالة

شهد العقد الماضي نموا لم يسبق له مثيل في كل من الطب الحيوي في إنتاج كمية من البيانات والكتابات المنشورة مناقشته.التقدم في الأساليب البيولوجية الحاسوبية وتغير ملحوظ في حجم بحوث الطب الإحيائي. الجينوم الكامل يمكن الآن يتسلسل في غضون أشهر وحتى أسابيع ، الأساليب الحسابية الإسراع في تحديد عشرات الآلاف من الجينات على نطاق واسع وأساليب تجريبية. البيانات المتولدة عن هذه التجارب هي مرتبطة بدرجه عالية ؛ من نتائج تحليل تسلسل والجزئي المصفوفات تعتمد على المعلومات الفنية ونقل الإشارة المذكورة في الممرات لاستعراض الأقران لمنشورات الادله. على الرغم من العلماء في هذا المجال ، يتم مساعدتهم بالعديد من قواعد البيانات على الانترنت من التفاعلات الكيميائية الحيوية ، لكن في الوقت الحاضر اغلبيه هذه هي المنسقة : العمل بشكل مكثف من قبل خبراء المجال. ولذلك فقد السعي لاستخراج المعلومات من النص بنشاط في محاولة لانتزاع المعرفة من المواد المنشورة والإسراع في عملية التدوين في قواعد البيانات إلى حد كبير. استخراج أداة من شأنه ليس فقط توفيرا للوقت والجهد، ولكنها أيضا تمهد الطريق لاكتشاف معلومات غير معروفة حتى الآن ضمنا في النص. هذه الفرضية تعرض النظام الذي يحمل اسم PIELG لاستخراج التفاعلات البروتين في الخلاصات الطبية الحيوية. ونهجنا يقوم على تقسيم الأولى الخلاصات إلى الجمل البسيطة. بعد ذلك ، التعليم البيولوجي الكيانات مع مساعدة الطبية واللغوية PIELG.

وأخيرا، والاستخراج وذلك من خلال تحليل التفاعلات الكامل المطابقة محتويات نحوي أدوار ومجموعات كبيرة لغويا. النظام يتناول الأحكام ومعقدة ومتداخلة متعددة مقتطفات التفاعلات المحددة في حكم صادر بحقه. ونحن في نطاق التجارب يقتصر على المستخلصات التي تصف حقوق وظيفة البروتين. فان مجموعة من النظام يتم اختيار لتقييم المقترح من البروتين التفاعل المصادقة طريقه. هذا هو مجموعة مختارة لتكون البروتينات في الوقت الراهن عن أن يكون في الأسنان تشكيل الأدوار والمشاركة في عملية تكوين الأسنان. نفذنا عمليات التقييم التجريبية للنظام PIELG.

التفاعلات المستخرجة من النظام تفحص يدويا للدقة والتذكير. حساسية هذا النظام هو الذي توليه يذكر التدبير، محسوبة على أنها النسبة بين التفاعلات المستخرجة بطريقة صحيحة والتفاعلات في هذا النص. الدقة هي مقياس لصحة النظام من خلال قياس عدد مرات النتائج المستخرجة بشكل صحيح بالمقارنة مع العدد الكلي للنتائج. ونحن تجريبي وتبين النتائج أن النظام المعروضة هنا يحقق أداء أفضل من دون الحاجة إلى إنشاء نمط من الدليل (من قبل المستخدم) التي تلزم لهذه النظم الأخرى. إن مجال تجاربنا محدود إلى الملخصات التي توصف البروتين الإنسانية. إن منظومة PIELG مختارة بأن تقيم طريقة تفاعل التي تعتبر الآن لها أدوارا في عملية تكوين الأسنان. أنجزت تقييم أنظمة PIELG تجريبية. التفاعلات المستخلصة بواسطة نظام PIELG مفحوصة بشكل يدوي. النظام أعطى حساسية حيث استخلصت النسبة بين التفاعلات بشكل صحيح و تقدم التفاعلات في النص. الدقة إن لم تكن النتائج استخلص بشكل صحيح المقارنة مع الرقم العام النتائج. تعرض نتائجنا التجريبية ذلك الذي نظام PIELG مقدم هنا أحرز أداء بشكل أفضل بدون الحاجة من خلق النموذج يدوي (بواسطة المستخدم) أي إن مطلوب لأجل هؤلاء الأنظمة الأخرى.