

### Yıldız Teknik Üniversitesi – Bilgisayar Mühendisliği Bölümü

**Akademik Personel Proje Öneri Formu**

**Danışmanın Adı Soyadı:** M. Özgür CİNGİZ **Proje Türü:** Bilgisayar Bitirme

|  |
| --- |
| **Projenin Konusu:** Moleküler Literatür Verilerinin Entegrasyonu |
| **Projenin Amacı:**  Normal bir hücrenin fizyolojisi, gelişmesi veya hastalığa dönüşmesi ve ilerlemesi durumunun daha iyi anlaşılabilmesi için tüm-genom seviyesinde genler arasındaki sebepsel (causal) etkileşim (interaction) mekanizmasının ortaya çıkarılması çok önemlidir . Bu alanda moleküler bazda ilişkileri farklı deneylerle elde eden çok çeşitli veri kümeleri bulunmaktadır. Bu veri kümelerinin en büyük problemi standart bir şekilde entegrasyonun yapılamamasıdır. Bu alanda protein-gen ilişkilerini veren veri kümelerinin birleştirilmesiyle farklı gen ağı çıkarım algoritmalarıyla elde edilen gen ilişkilerinin doğrulaması daha başarılı bir şekilde yapılabilecektir.  Çalışmamız kapsamında yolak (pathway) veri kümelerinden elde edilen gen-protein ilişki veri kümeleri ve direk gen-protein ilişkilerini veren Biogrid, Intact, Mint, Innatedb gibi veri kümelerindeki tüm ilişkiler birleştirilmesi hedeflenmektedir. Bu veri kümelerinin kullandığı belirteç (ID) farklılık göstermekte ve ilgili veri kümelerindeki gen isimlerinin birden fazla farklı gösterimi bulunabilmektedir. Bu ve bunun gibi veri kümelerine özgü problemlere çözüm getirilmeye çalışılacaktır. Bunun dışında farklı veri kümelerinde yer alan ortak moleküler ilişkiler Venn şeması halinde sunulması hedeflenmektedir.  Çalışmanın ikinci amacı ise deney tiplerine göre yönlü ve yönsüz olarak belirlenmiş moleküler ilişkilerin belirlenmesidir. Buna göre ilgili gen listeleri kullanıcıdan elde edilerek olası ilişkiler çıkartılacak ve ilgili görselleştirme paketiyle gen ağı çizilecektir. Bunun dışında farklı canlılara ve çeşitli deney tiplerine göre kullanıcılar program içerisinde filter uygulayabilecek böylece uygulamanın fonksiyonelliği arttırılacaktır. |
| **Projenin Çıktıları:**  Proje R paketi olarak tasarlanmakla birlikte R paketiyle birlikte web ara yüzünün tasarlanması hedeflenmektedir. Buna göre kullanıcılar gen/protein isimlerini sisteme girebilmekte ve hedef genle ilgili moleküler ilişkiler kullanıcıya sunulabilmektedir. Bunun dışında kullanıcılara organizma bazlı filtreleme, deney tiplerine göre filtreleme seçenekleri sunulması planlamakta ve programın işlevselliği arttırılması hedeflenmektedir. Moleküler ilişkiler aynı zamanda görselleştirilerek kullanıcılara sunulacaktır. |
| **Önerilen Kaynaklar:**   * Mora, Antonio, and Ian M. Donaldson. "iRefR: an R package to manipulate the iRefIndex consolidated protein interaction database." BMC bioinformatics 12.1 (2011): 455. * Altay, Gökmen, Nejla Altay, and David Neal. "Global assessment of network inference algorithms based on available literature of gene/protein interactions."Turkish Journal of Biology 37.5 (2013): 547-555. * Olsen, Catharina, et al. "Relevance of different prior knowledge sources for inferring gene interaction networks." Frontiers in genetics 5 (2014). |
| **Projenin Özellikleri:**  Proje ekibinin boyutu:  1 öğrenci  2 öğrenci  Rekabet edebilecek ekip sayısı:  1 ekip  2 ekip  3 ekip |