



BİYOWEB: İNTERNET TABANLI BİYOENFORMATİK ARAYÜZÜ

Yunus ÖZCAN, Murat ERTEKİN, Gıyasettin ÖZCAN

Dumlupınar Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Kütahya,
yunus.ozcan@ogr.dpu.edu.tr, murat.ertekin@ogr.dpu.edu.tr, giyasettin.ozcan@dpu.edu.tr

ÖZET

Bu çalışmada biyoenformatik alanında kullanılan algoritmalar, bir web uygulaması olarak geliştirilmiştir. Özellikle biyoloji, genetik ve bilgi erişim sistemi çalışan öğrencilerin bu algoritmaları tanıması, nasıl çalıştığı hakkında bilgi vermesi ve bu alandaki örnek programların çalıştırması temelinde kurulmuştur. Amacı üniversitemizin bu alanda yaptığı çalışmalardan başka kişilerinde yararlanmasını sağlamak ve üniversitemizi bu alanda tanıtımını yapmaktır. PHP ile yazılan web sayfalarında c++ kodları çalıştırılmış ve sonuçlar javascript yardımıyla kullanıcıya geri döndürülmüştür.

Anahtar Kelimeler: *Biyoenformatik, php arayüzü, mySQL, c++ uygulamaları, sekans hizalama.*

BIOWEB:INTERNET BASED BIOINFORMATİC INTERFACE

ABSTRACT

In this work, we design a web application that are used in bioinformatics field. Goal of the work is a toolbox that helps to the students in biology, genetics and information retrieval systems, so that they can understand related algorithms. As a consequence of this application, we aim that members of other institutions make use of the application which could improve credibility of our univesity. Inside Php code snippets, we execute C++ programs which execute bioinformatics algorithms. The results of the algorithms are returned back with the constribution of javascript code blocks .

Keywords: *Bioinformatics, php interface, mySQL, c++ applications, sequence alignment.*

1. GİRİŞ

Biyoloji veyahut genetik veri analiz yöntemlerinde bilgisayar kullanılması kaçınılmaz bir hale gelmiştir. Zira elde edilen veri miktarını insanoğlunun elle yapması imkânsızdır. Öte yandan elde edilen veri kütlesi o kadar büyük ve o kadar karmaşıktır ki insanoğlunun beklediği cevapları temin edecek hesap gücü maalesef bilgisayarlar tarafından dahi sunulmamaktadır.

Biyoinformatik açısından bakacak olursak DNA, RNA sekansları arasında benzerlik ya ortak desenlerin bulunması problemlerden sadece birisidir[1]. Öte yandan çok sayıda DNA-RNA sekansının benzerlik amacı ile hizalanması prosedürü oldukça karmaşıktır. DNA'nın atadan yeni nesille geçişi esnasında sık mutasyona uğramış olması problemi daha da karmaşık hale getirmektedir.

Smith-Waterman algoritması[2], Burrows-Wheeler ile arama algoritması[3] gibi yöntemler yeni nesil hizalama tekniklerine ön ayak olmuşlardır. Öte yandan Gibbs Örnekleme[4], Saklı Markov[5] Modelleri, Rastgele Projeksiyon[6], gibi yöntemler hizalama konusunda olasılıksal bakış açısı sunmaktadır.

Bir diğer önemli bir konu ise Filojeni'dir. Bu alan genlerin ortak atasını bulmaya çalışmaktadır. Mutasyonlara rağmen ortak atayı tahmin etmek zor olsa da bir çok hastalığın tedavisinde önemli yer alabilecek konumdadır.

DNA, yada RNA'nın bu kadar önemli olmasının nedeni protein içinde yer alan atomların dizilimini belirlemesinden kaynaklanmaktadır. Dolayısıyla DNA bilgisine bağlı olarak protein içinde yer alan atomların hangi şekilde üç boyutlu olarak bağlanacağını tahmin etmek bilgisayarların dahi kapasitesini aşmaktadır.

Yukarıda bahsedilen problemlere çözüm bulunabilmesi için alanı çok iyi anlayan öğrencilere ihtiyaç vardır. Bu amaçla web tabanlı arayüz tasarımlarının geliştirilmesi elzem olmuştur.

2. MATERİYAL VE METOT

Bu alanda yazılmış ve ya okulun öğrencileri tarafından yazılan c++ uygulamaları gerekli düzenlemeler yapıldıktan sonra PHP sayfasından girdi değeri alması sağlanmıştır. C++ uygulaması bu girdi değerleriyle çalıştırılarak kullanımcıya sonuç döndürülmüştür. Uygulamada C++ ve PHP seçilmesinin sebebi platformdan bağımsız şekilde tüm işletim sistemlerinde aynı sonucu verebilmesi için düşünülmüştür.

Kullanılan programlar:

C++ Editörü, PHP Editörü, Javascript, HTML/CSS, MYSQL, WampServer şeklindedir.

İnternet sitesinde taslak olarak; Üst sol kısımda üniversitemizin logosu ve bölüm ve dersin adı belirtilmiştir. Üst sağ kısımda üniversitemizin resmi bulunmaktadır. Bu kısmın hemen altında menü kısmı yer almaktadır.

Orta kısmın arka planında üniversitemizin bir başka görseline yer verilmiştir. En alt kısımda da özlük hakları belirtilmiştir. Menü kısmının yanında akıllı arama bulunmaktadır.

Ana sayfa iki kısımdan oluşmaktadır. Sol kısımda algoritmaların belirli kriterlere(en çok okunan-ekleme tarihine göre) göre listelendiği bir bölüm bulunmaktadır. Bu işlem veri tabanına bağlantı yapılarak gerçekleştirilmiştir. Kullanıcı bu menüdeki algoritmaya tıkladığında link içerisine algoritmanın id numarası gömülerek algoritmalar sayfasına yönlendirmektedir.

Sağ kısımda biyoinformatik hakkında görsel verilerin olduğu bir resim galerisi bulunmaktadır. Resim galerisinin altında biyoinformatik hakkında genel bilgi verilmiştir.

3.UYGULAMA

Web uygulaması, birçok ara yüzünden oluşmaktadır. Zira yapılacak C++ uygulamalarının her birisi başka bir grup tarafından hazırlanıp sisteme entegre edilmesi planlanmıştır.

3.1. Kod Yükleme Sayfası

Sisteme yüklenen algoritmaların adları listelenmiştir. Bu işlem veri tabanına bağlantı yapılarak gerçekleştirilmiştir. Kullanıcı bu menüdeki algoritmaya tıkladığında link içerisine algoritmanın id numarası gömülerek algoritmalar sayfasına yönlendirmektedir.

Algoritmalar Sayfası;

Link ile bu sayfaya yönlendirilen algoritmalara göre sayfa tasarlanmaktadır.

Sayfa iki kısımdan oluşmaktadır.

Sol kısımda algoritmaların belirli kriterlere(en çok okunan-ekleme tarihine göre) göre listelendiği bir bölüm bulunmaktadır. Bu işlem veri tabanına bağlantı yapılarak gerçekleştirilmiştir. Kullanıcı bu menüdeki algoritmaya tıkladığında link içerisine algoritmanın id numarası gömülerek algoritmalar sayfasına yönlendirmektedir.

Sağ kısımda ilk olarak açıklama kısmının olduğu alan gelmektedir. Bu kısım link ile gelen algoritma id numarasına göre doldurulmaktadır. Bu kısımda algoritmanın tanıtımı, çalışma mantığı, kullanıldığı alanlar hakkında bilgi verilmiştir. İkinci olarak uygulama kısmının olduğu alan gelmektedir. Bu kısımda çalışan algoritma sayfadan girdi değerlerini alır. Alınan değer başka bir sayfaya yönlendirilir. Yönlendirilen sayfada "passthru" komutu yardımıyla program çalıştırılıp javascript ve ajax yardımıyla sonucu algoritmalar sayfasına post edilmeden yazdırılmaktadır. Bu kısmın hemen altında bu konu hakkında yazılan bilgilerin indirilmesinin sağlandığı link bulunmaktadır.

3.2. İletişim Sayfası

İletişim Sayfası;

Sayfada üniversitemizin ve bölümün iletişim bilgileri yer almaktadır.

3.3 Giriş Sayfası

Kullanıcının kullanıcı adı ve şifresinin kontrolünü yapıp uyarı geri dönüşleri yaparak doğru olduğunda kullanıcı paneline yönlendirmektedir.

3.4 Kullanıcı Paneli Sayfası

Sayfa iki kullanıcı yetkisine göre tasarlanmaktadır.

Birinci kullanıcı türünün erişimleri;

- Algoritma Ekleme
- Algoritmaların herhangi birinde düzenleme yapabilme ve ya silme
- Eklenen algoritmaların bilgilerine ulaşabilme
- Kullanıcı şifresini değiştirme

İkinci kullanıcı türünün erişimleri;

- Algoritma Ekleme
- Eklediği algoritmayı ekleme veya silme
- Kullanıcı şifresini değiştirme
- Kullanıcı adı ve şifresinden gelen bu bilgilerle iki taslaktan biri yüklenmektedir.
- Algoritmanın ekleneceği alanda;

Algoritma adı, algoritma hakkında bilgi, uygulama dosyası, kullanıcıların bu dosyaları indirebilmeleri için zip dosyası, uygulamada girilecek girdi sayısı, girilen girdi sayısına göre yeni açılan alanda bu girdilerin neler olduğu ve türü, son olarak yükleme yaptığını kabul eden bir kabul formu bulunmaktadır. Kullanıcıların nasıl yükleme yapması gerektiğini program dosyasının nasıl düzenleneceğini anlatan bir Word dosyası bulunmaktadır.

Kullanıcıların girdi değerleri alınıp başka sayfada kontrol edilip bilgilendirilmesi yapılacak bir hata ile karşılaşmadığı takdirde veri tabanına kayıt yapması sağlanacaktır. Kullanıcı bu konuda bilgilendirilecektir.

Algoritma düzenleme sayfasında birinci kullanıcı istediği algorithmada düzenleme yapması ve silmesi sağlanmıştır. Düzenleme yapılacağı algoritma seçilip algoritma ekleme sayfasına benzeyen sayfa doldurulacaktır. İkinci kullanıcı türüyle giren kişiler sadece kendi eklediği algoritmalarından birini seçip düzenleme için sayfayı dolduracaktır. Aynı şekilde bu girdi değerlerinin işlemleri başka sayfada yapıp hata ile karşılaşmadığı takdirde veri tabanında güncellenecektir. Kullanıcı bu konuda bilgilendirilecektir. Şifre değiştirme ekranında eski şifresi istenilip doğru girdiği takdirde yeni şifresiyle değiştirilmesi sağlanacaktır. Bu işlem için kullanıcı bilgilendirilecektir.

Sadece birinci kullanıcı için yazılan algoritma ları göster sayfası seçilen algoritmanın kim tarafından hangi tarihte yüklendiğini hangi dosyaları kullandığını göstermektedir.

4.SONUÇLAR VE TARTIŞMA

Biyoinformatik hakkında Türkçe bir kaynak sağlanmış olup kullanıcıların bu kaynaklara ulaşması sağlanmıştır. Biyoinformatik adına Türkçe bir veri kümesi oluşturulmuştur. Bu konuyla ilgilenenlerin kaynaklara rahat bir şekilde erişmesine imkân sağlanmıştır.

KAYNAKLAR

- [1] R. Durbin, S. Eddy, A. Krogh, and G. Mitchison, "Biological Sequence Analysis, Cambridge University Press, Cambridge, (1998).
- [2] T.F. Smith, and M.S. Waterman, "Identification of Common Molecular Subsequences", Journal of Molecular Biology, 147: 195–197, (1981).
- [3] P. Ferragina, and G. Manzini, "Opportunistic data structures with applications", In 41st Annual Symposium on Foundations of Computer Science Proceedings, 390–398. IEEE, (2000).
- [4] S. Geman, D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images". IEEE Transactions on Pattern Analysis and Machine Intelligence 6 (6): 721–741, (1984).
- [5] L. Rabiner "A tutorial on Hidden Markov Models and selected applications in speech recognition", Proceedings of the IEEE 77 (2): 257–286, (1989).
- [6] J. Buhler, M. Tompa, "Finding motifs using random projections", Proceeding of RECOMB '01 Proceedings of the fifth annual international conference on Computational biology, ACM New York, NY, USA (2001).