

T.C DOKUZ EYLÜL ÜNİVERSİTESİ FEN FAKÜLTESİ BİLGİSAYAR BİLİMLERİ BÖLÜMÜ

SESLİ KOMUT UYGULAMASI

Burak GÖMEÇ

Ağustos, 2020

İZMİR

ÖZET

Proje, son zamanlarda oldukça popüler olan sesli komut üzerine inşa edilmiştir.

Sesli komut günümüzde mobil işletim sistemlerinde, kişisel ev asistanlarında ve

bankacılık sektörü gibi birçok alanda oldukça sık kullanılmaktadır.

Uygulama Visual Studio üzerinde Windows Form Uygulamasında Türkçe

kütüphane yetersizliğinden dolayı İngilizce olarak geliştirilmiştir. Uygulamanın başarı

oranı mikrofon, telaffuz seviyesi gibi etkenlere bağlıdır, başarı oranını arttırmak için

yapılanlar raporda detaylıca açıklanmıştır. Uygulamanın ilk kısmında kullanıcı

tarafından sesli komutlar ile belirli işlemler yapabilmektedir. Uygulamanın ikinci

kısmında ise sqlite veri tabanı kullanılarak hem sesli komut hem de klavyeden giriş

yapılarak alınan veriler, veri tabanına kayıt edilerek uygulamada güncel olarak

gösterilmektedir.

Raporda hedeflenen temel kazanımlardan biri sesli asistanların temelinde bir karar

yapısı yer aldığı ve elimizde büyük bir veri var olduğu durumda, istenilen büyüklükte

bir uygulama geliştirilebileceğidir.

Anahtar Sözcükler: Sesli komut, Windows form uygulaması, Sqlite veri tabanı

Speech sınıfı, Yapay zekâ

2

ABSTRACT

The project is built on voice command, which has been hugely popular recently. Voice command is used quite frequently today in mobile operating systems, personal home assistants, and many other areas such as the banking sector.

The application was developed in English language due to the Turkish language library insufficiency in Windows Form Application on Visual Studio. The success rate of the application depends on factors such as microphone and pronunciation level. The actions to increase the success rate are explained in detail in the report. In the first part of the application, the user can perform certain operations with voice commands. In the second part of the application, the data received by using both the voice command and the keyboard by using the sqlite database are recorded in the database and displayed in the application.

One of the main gains targeted in the report is that when voice assistants are based on a decision structure and we have a big data, an application of the desired size can be developed.

Keywords: Voice commands, Windows form application, Sqlite database, Speech class, Artifical intelligence

İÇİNDEKİLER

ÖZET	2
ABSTRACT	3
İÇİNDEKİLER	4
BÖLÜM BİR-GİRİŞ	5
1.1. Projenin konusu	6
1.2. Uygulamanın Algoritması	7
BÖLÜM İKİ-UYGULAMA İÇERİĞİ VE DEĞERLENDİRME	9
2.1. Uygulama ve Ekran Görüntüleri	9
2.1.1 Personal Voice Assistant	9
2.1.2. Products	16
2.2 SONUÇ VE DEĞERLENDİRME	21
KAYNAKCA	22

BÖLÜM BİR

GİRİŞ

Bu projede ele alınan sesli komut temellerine üzerine inşa edilen bu uygulamayı seçmeninin sebebi; son zamanlarda oldukça popüler olan sesli asistanların temelini ve çalışma algoritmasını ortaya koymaktır.

Uygulamayla ilgili Türkçe kütüphane .NET platformunda olmadığından dolayı dil İngilizce olarak seçilmiştir. Türkçe kullanım için ücretli kaynaklar yerine Microsoft'un "speech" kütüphanesi kullanılmıştır.

Uygulamanın başarı oranını düşüren bazı unsurlar şunlardır; mikrofonun kalitesi, asistanın verdiği cevapları mikrofon tarafından algılanıp tekrar komut vermesi ve telaffuz hatalarından kaynaklı bazı yanlış veri girişleridir. Bunun için uygulama belirli kelimelere ve cümlelere karşı hassas şekilde tasarlanıp geliştirilmiştir ve bu şekilde ele alınmasının faydası olarak her kullanıcı için telaffuz seviyesi hangi durumda olursa olsun yüksek oranda başarı sağlayarak programa tanıtılan hassas kelimeleri anlayabilmesi sağlanmıştır.

Ünlü şirketlerin kullandığı yapay zekâ başlığı altında sesli asistanlarında bir çok alanda olduğu gibi big data kullanıldığı bariz şekilde ortaya çıkmıştır. Bu uygulama bazında elimizde ne kadar veri var ise uygulamanın boyutu ve başarı oranı o kadar artmaktadır. Sesli asistanların arka planda bir karar mekanizması olduğu ve söylenen komuta karşılık arka planda ona tanımlanan bir komutu, yapılması beklenen görevi yaptıkları elde edilmiştir.

Uygulama bilgisayarda İngilizce dil paketi istediği için test edilen bilgisayarda dil paketi eksik olduğu durumda kullanıcı bir mesaj ile uyarılmaktadır. Bu yüzden projenin olduğu ana klasöre Microsoft'un dil paketinin kurulum dosyası eklenerek olası bir dil paketi hatasının önüne geçilmiştir.

Uygulama geliştirilirken aktif olarak Github üzerine yüklemeler yapılmıştır.

1.1. Projenin konusu

Projenin konusu temel olarak sesli komutlar kullanılarak uygulama tarafından belirlenen işlemler yapılması istenilmektedir.

Uygulama sesli asistan ve bir ürün-stok sistemi olarak 2 kısımdan oluşmaktadır. Uygulamanın algoritması 2.2. Uygulamanın Algoritması kısmında, ekran görüntüleriyle birlikte açıklamalar sunulduğu kısım 2.3 Uygulama ve Ekran Görüntüleri kısmında örneklerle detaylıca açıklanmıştır.

Sesli asistanlarla ilgili merak edilenler, Türkçe bir sesli asistan hangi dilde geliştirilmeli, sesli komutların temelinde neler var olduğu gibi sorular rapor boyunca açıklanmaya çalışılmıştır.

1.2. Uygulamanın Algoritması

Uygulama Visual Studio üzerinde C# dili ile geliştirilmiştir. Uygulamada kullanılan komutları sese, sesi de yazıya dönüştüren sınıf "speech" tir.

Uygulama 2 adet Form içermektedir. İlk formda (personal voice asistant) 11 adet farklı komut sesli söylenerek komutlar karşılığında cevaplar sesli olarak gelmektedir ve uygulama tarafından belirli işlemler yapılmaktadır. Komutlar şu şekildedir; "Hello", "Open Paint", "Open Google", "Open Youtube", "What Time Is It", "How Are You", "Hey Assintant", "Exit The Application", "Stop Listen", "Open Other Form", "Show Todays Exchange Rate".

Formda mikrofon simgesine tıklayarak, uygulama sizi dinlemeye başlar ve sesli olarak bildirim vermektedir. Örneğin, uygulamaya "show todays exchange rate" şeklinde bir komut söylendiğinde o güne ait güncel Euro ve Dolar kuru bilgileri ekrana gelmektedir. (Bu veriler http://www.tcmb.gov.tr/kurlar/today.xml sayfasından ele alınmıştır.)

Uygulama ilk kısmında 11 adet komutu karşılaştırma yaparak kontrol etmektedir. Program 11 adet komuta hassas olduğundan dolayı başka bir komut söylendiğinde cevap vermemekte veya 11 komutun içinden telaffuzuna benzettiği komutu seçmektedir. Uygulama belirli komutlara hassas şekilde tasarlanmasaydı test edilen duruma göre telaffuzdan kaynaklı başarı oranı çok azalmaktadır. Bu yüzden belirli komutlara hassasiyet sağlanarak başarı oranı arttırılmıştır.

İkinci Formda(products) ise klavyeden veri girişi ve sesli komut ile bir ürün-stok uygulaması tasarlanmıştır. Veri tabanı olarak hızından ve kullanım kolaylığından sqlite veri tabanı kullanılmıştır.

Formda ürüne ait isim, marka, adet, fiyat olarak 4 adet veri alınmaktadır. Bu veriler isteğe bağlı olarak hem sesli olarak hem de klavyeden alınarak veri tabanına kayıt edilmektedir. Uygulama çalıştığı anda ve yeni ürün eklendiği anda tüm kayıt tablosu güncellenerek tüm ürünler ve bilgiler ekrana gelmektedir.

Uygulamanın başarı oranını arttırmak için ilk formda olduğu gibi ikinci formda da belirli markalara, ürünlere ve sayılara karşı bir hassasiyet sağlanmıştır. Örneğin uygulama adet ve fiyat bilgisi için 0-1000 arası sayılara, belirli markalara, belirli ürünlere hassasiyet sağlanmıştır. Buradaki temel sebep, uygulamayı test eden kişilerin telaffuz seviyesinden bağımsız olarak başarıyla test edebilmesi amaçlanmıştır.

BÖLÜM İKİ

UYGULAMA İÇERİĞİ VE DEĞERLENDİRME

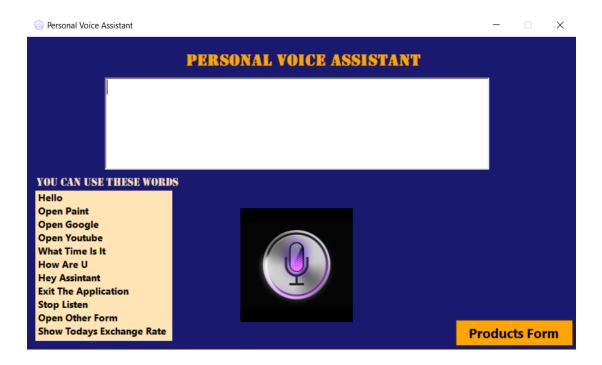
2.1. Uygulama ve Ekran Görüntüleri

Uygulama 2 kısımdan oluşmaktadır.

- Personal Voice Assistant
- Products

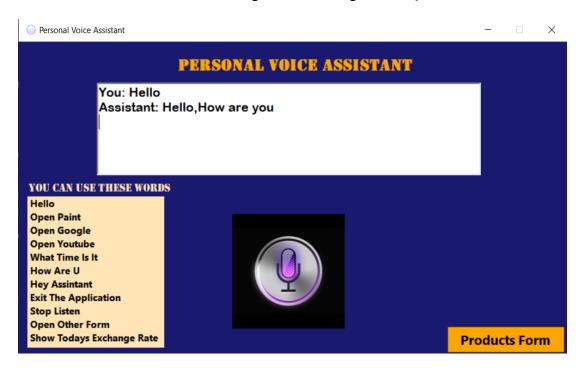
2.1.1 Personal Voice Assistant

Bu kısımda uygulamanın ilk formu ve tüm komutlar gösterilmektedir. İlk formun ekran görüntüsü aşağıda yer almaktadır. Uygulamayı başlatmak için mikrofon simgesine tıklanır ve solda tanımlı olan komutlardan birisi söylenerek uygulama çalışmaya başlar.



Ekran Görüntüsü 2.1.1 Personal Voice Assintant Formu

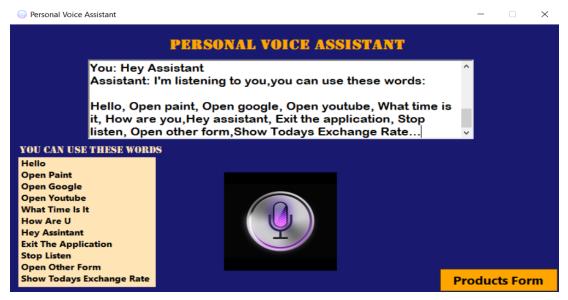
İlk olarak "Hello" komutu verildiğindeki durum gösterilmiştir.



Ekran Görüntüsü 2.1.1 "Hello" Komutu

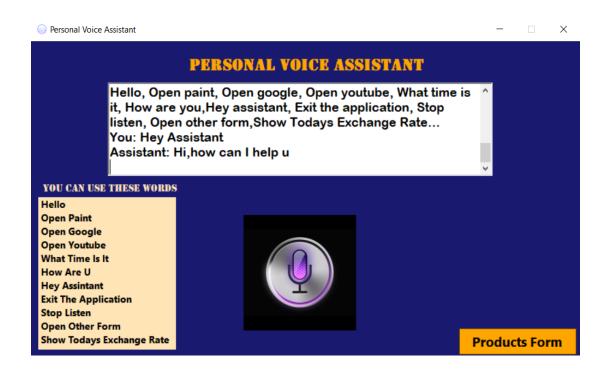
Uygulama "Hello" komutuna karşılık "Hello, How Are You" komutu vermektedir. Uygulama kapatılmadığı sürece veya "Stop Listen" komutu verilmediği sürece sürekli olarak kullanıcıyı dinlemektedir, bu defalarca aynı butona basılmasını önlemek ve daha verimli bellek kullanımı açısından yapılmıştır.

Uygulamaya "Hey Assintant" komutu verildiği durumu gösterilmiştir.



Ekran Görüntüsü 2.1.1 "Hey Assistant" Komutu

"Hey Assistant" komutu verildiğinde uygulama size kullanabileceğiniz komutları söylemektedir ve belirlenen iki cümleden biri olan "Hi, how can I help you" komutunu söylemektedir. Bu programın o andaki arka planda rastgele ürettiği sayıya bağlıdır. Arka planda üretilen sayı 1 ise ilk cümleyi 2 ise diğer cümleyi komut olarak alıp söylemektedir. Burada fark edilen durum mobil cihazlarda ve genel olarak sesli asistanların örnek olarak "Nasılsın" komutuna karşılık belirli 5-6 farklı komutu döndürmektedirler. Çıkarılan sonuç ise sesli asistanların büyük veri (big data) kullandıkları görülmüştür.



Ekran Görüntüsü 2.1.1 "Hey Assistant" Komutundaki Farklı Cevap

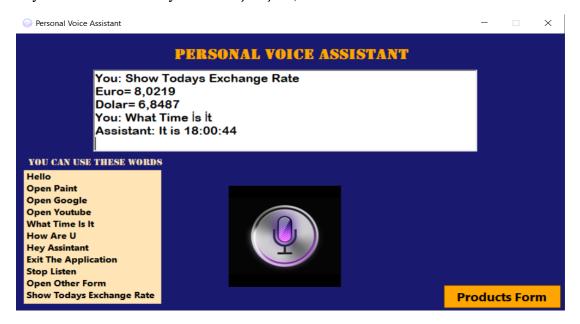
Uygulamaya "Open Paint" komutu verildiğinde Paint uygulamasını, "Open Google" komutunda varsayılan tarayıcıda Google ana sayfasını, "Open Youtube" komutunda ise tarayıcıda Youtube ana sayfasını açmaktadır.

"Show Todays Exchange Rate" komutu verildiğinde uygulama tarafından "https://www.tcmb.gov.tr/kurlar/today.xml" sayfasından güncel Euro ve Dolar kuru bilgileri çekilerek kullanıcıya gösterilmektedir. Verilerin listelenmesinde internet hızına bağlı olarak anlık gecikmeler yaşanabilmektedir.



Ekran Görüntüsü 2.1.1 "Show Todays Exchange Rate" Komutu

"What Time Is It" komutunda anlık olarak saat, dakika ve saniye bilgilerini kullanıcıya sesli olarak söylenip gösterilmektedir. (Her komutun baş harflerini büyük yapmak amacı ile "Globalization" sınıfı altında "ToTitleCase" metodu kullanılmıştır, büyük "i" kullanımı bu yüzden oluşmuştur.)



Ekran Görüntüsü 2.1.1 "What Time Is It" Komutu

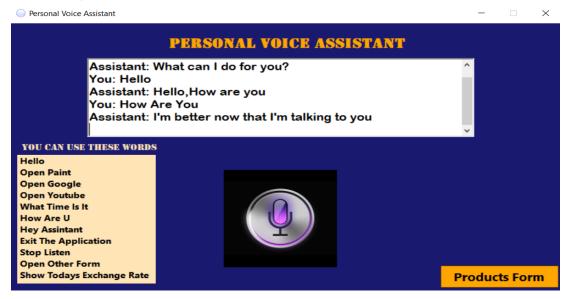
Uygulama "Exit The Application" komutu aldığında uygulamayı sonlandırmaktadır. "Open Other Form" komutunu aldığında ise ikinci form olan Products Formunu açmaktadır.

Uygulamanın kullanıcıyı dinlemeyi bırakması için "Stop Listen" komutu tanımlanmıştır. Bu komut söylendikten sonra mikrofon simgesinin üzerinde "Click to Active" şeklinde bir yazı belirmektedir, mikrofon simgesine tekrar tıklanarak uygulama aktif hale gelir ve kullanıcıyı dinlemeye devam etmektedir.



Ekran Görüntüsü 2.1.1 "Stop Listen" Komutu

Son olarak "How Are You" komutu ve karşılığı gösterilmiştir. Bazı durumlarda telaffuzdan dolayı uygulama "How Are You" komutunu "Hello" komutu olarak ele almaktadır, bu durumda komut net bir telaffuzla tekrar edilmelidir.



Ekran Görüntüsü 2.1.1 "How Are You" Komutu

2.1.2. Products

Bu kısımda uygulamanın ikinci kısmı olan Products Formu tanıtılacaktır. Uygulamanın bu kısmında bir ürün-stok sistemi ile sesli komutlar birleştirilmiştir. Uygulamada veri tabanı olarak hız ve kullanım kolaylığı açısından Sqlite veri tabanı kullanılmıştır.



Ekran Görüntüsü 2.1.2 Ürünler Formu

Formda yer aldığı şekilde ürünün adı, markası, fiyatı ve adedi klavyeden giriş yapıldığında "Add Manually" butonuna tıklanarak bilgiler elle veri tabanına kayıt olmaktadır. Kullanıcı kayıt işlemini sesli komut ile gerçekleştirmek istediğinde ise "Start Voice Control" butonuna tıklar ve uygulama tamamen sesli bir yapıya geçmektedir. Product List başlığı altında yer alan tabloda güncel olarak veri tabanından ürün ve ürünün detayları gelmektedir, bir ürün eklendiğinde ve uygulama yeniden başlatıldığında tablo kendini sürekli güncellemektedir.

İlk olarak klavyeden veri girişi yapıldığı durum ele alınmıştır.



Ekran Görüntüsü 2.1.2 Ürünleri Elle Ekleme Durumu

Örnek olarak ürün ismine kulaklık, markasına "Haylou", fiyatına 50 birim, adedine 150 girilmiştir. "Add Manually" butonuna tıklanarak girilen ürün detayları veri tabanına eklenmektedir ve tablo güncellenmektedir. Veri tabanının güncel hali aşağıda gösterilmiştir.



Ekran Görüntüsü 2.1.2 Ürünlerin Klavye İle Girilmesi Durumu

Yukarıdaki ekran görüntüsünde görüldüğü gibi ürün ve ürünün detayları veri tabanına kayıt olmuştur. Uygulama veri tabanından bilgileri alarak tabloda kullanıcıya göstermektedir ve ürünün ismi, markası, fiyatı, adedi kısımları tekrar boş bir alan haline gelmektedir.

Bir diğer kısım olarak sesli komutlar ile ürün ekleme durumu ele alınmıştır. Uygulamanın ilk Formunda olduğu gibi Products Formunda da başarı oranını arttırmak için belirli ürünlere, markalara ve sayılara karşı hassasiyet sağlanmıştır. Kullanıcı tarafından kullanılabilecek ürünler ve markalar "You Can Use These Words" başlığı altındaki kısımda belirtilmiştir.

"Start Voice Control" butonuna tıklandığı anda sesli komut algoritması arka planda çalışmaya başlamaktadır ve butonun ismi değişerek "Say The Commands" şeklini almaktadır, böylece kullanıcıya komut vermesi gerektiği belirtilmiştir.

"Name" şeklinde bir komut verdiğimizde uygulama name kısmına odaklanıp giriş alanına yeşil arka plan vermektedir. Uygulama name kısmına odaklandığı anda örneğin, "Music Player" şeklinde bir komut verildiğinde telaffuz başarılı bir şekilde söylenirse uygulama name kısmına "Music Player" yazmaktadır.



Ekran Görüntüsü 2.1.2 "Name" Komutu



Ekran Görüntüsü 2.1.2 "Music Player"

Yukarıdaki ekran görüntüsünde sesli olarak "Music Player" komutu söylenmiştir.

Kullanıcı ürüne bir marka girmek için "Brand" komutu vererek uygulamanın marka kısmına odaklanması sağlanır ve sağ kısımda yer alan markalardan biri söylenir. "Brand" i telaffuz ederken bazı durumlarda uygulama söyleneni anlamayabilmektedir bu durumda doğru bir telaffuzla tekrar komut verilmelidir.

Uygulamaya "Brand" komutu verildikten marka olarak "Lg" markası sesli olarak söylenmiştir.



Ekran Görüntüsü 2.1.2 "Lg"

Eğer sesli olarak söylenen ürün ismi, markası, fiyatı, adedi bilgileri uygulama tarafından yanlış anlaşıldığı durumda değiştirilmek istenen alan hangisi ise örneğin ürünün markası ise "Brand" komutu verilerek uygulama tekrar marka kısmına odaklanır ve yeni söylenen markayı alır.

Ürüne ait bir fiyat girmek için "Price" komutu verilerek 0-1000 arasında bir sayı söylenmelidir, bazı durumlarda yine telaffuzdan kaynaklı istenen sayı yerine başka bir sayı gelebilmektedir, böyle bir durumda girilmek istenen sayı tekrar edilmelidir. Aynı şekilde ürüne adet girmek için "Piece" komutu verilerek yine 0-1000 arasında bir sayı söylenmelidir.



Ekran Görüntüsü 2.1.2 Bir Ürünün Veri Tabanına Kaydı

Yukarıdaki ekran görüntüsünde "Music Player" a fiyat olarak 320 birim, adet olarak ise 8 verilmiştir. Ürüne ait tüm alanlar dolduğunda uygulama tarafından ürün veri tabanına kayıt edilir ve anlık olarak tablo güncellenerek kullanıcıya son eklediği ürün gösterilmektedir.

2.2 SONUÇ VE DEĞERLENDİRME

Sonuç olarak projede ele alınan sesli komut uygulamasıyla, günümüzde kullanılan sesli ev asistanları gibi bir çok sesli komutla çalışan uygulamanın temelinde bir karar yapısına sahip olduğu, söylenen komuta karşılık arka planda bir karşılaştırma yaptıkları ve elimizde ne kadar çok büyük veri var ise bu verilerle uygulama eğitilerek sistemin başarı oranının o kadar arttığı ve sistemin kullanım alanının o kadar büyüdüğü sonuçları ortaya çıkmıştır.

Uygulamanın olumlu sonucu programı test eden kişinin kelimeleri telaffuz etme seviyesi ve mikrofon durumunu ele almadığımızda istenen her komutu bekleme süresi olmadan gerçekleştirmesidir. Olumsuz sonucu en kötü durum olarak dış sesi çok alan bir mikrofon kullanılması, test edilen bilgisayarın dahili mikrofonu kullanıldığında bazen asistanın kendi verdiği cevapları da alması ve test eden kişinin kelimeleri yanlış telaffuz etmesi durumunda uygulamanın başarı oranını düşmesidir. Diğer bir olumsuz sonuç ise uygulamanın test edildiği bilgisayarda İngilizce ses kütüphanesi olmadığı durumdur bu sebepten dolayı projenin olduğu ana klasöre Microsoft'un İngilizce ses kütüphanesinin kurulum dosyası eklenerek bu sorun ortadan kaldırılmıştır.

Visual Studio üzerinde C# diliyle geliştirilen Windows Form uygulamasında Türkçe kütüphane olmadığından dolayı uygulamanın başarı oranı Türkçe bir asistana göre bazı İngilizce kelimelerin telaffuzundan dolayı daha düşüktür. Bu sebeple ileride başarı oranı daha yüksek Türkçe bir sesli asistan yapılmak istendiğinde Python ile geliştirilme yapılabilir ya da ücretli olarak sunulan sesli kaynaklar kullanılabilir.

KAYNAKÇA

WEB_1. (2020). Microsoft web site. https://docs.microsoft.com/tr-tr/dotnet/api/system.speech.synthesis?view=netframework-4.8, 05/07/2020

WEB_2. (2020). Microsoft web site. https://docs.microsoft.com/tr-tr/dotnet/api/system.speech.recognition?view=netframework-4.8, 05/07/2020

WEB_3. (2020). Microsoft web site. https://docs.microsoft.com/tr-tr/dotnet/api/system.speech.recognition.choices?view=netframework-4.8, 07/07/2020

WEB_4. (2020). Microsoft web site. https://docs.microsoft.com/tr-tr/dotnet/api/system.speech.recognition.grammar?view=netframework-4.8, 07/07/2020

WEB_5. (2020). Microsoft web site. https://docs.microsoft.com/tr-tr/dotnet/api/system.speech.recognition.recognitioneventargs?view=netframework-4.8, 08/07/2020