

راهنمای استفاده از واسط کاربردی پیکره احساس سنتی پرس

فرمت ذخیره سازی اسناد در پیکره احساس سنتی پرس در قالب ایکس.ام.ال^۱ است. هرچند با استفاده از فرمت حاضر نیز می توان اطلاعات مورد نظر را از اسناد پیکره استخراج نمود، جهت سهولت در استفاده از پیکره، یک واسط کاربردی در قالب یک فایل دی.ال.ال^۲ مهیا شده و اطلاعات پایه و اساسی را از طریق این واسط کاربردی می توان از اسناد پیکره استخراج نمود (به زودی کتابخانه معادل برای زبان جاوا نیز پیاده سازی خواهد شد). در اینجا به شرح وظایف توابع موجود در این واسط کاربردی پرداخته و چگونگی استفاده از آنها را شرح خواهیم داد. **توصیه می شود که پیش از استفاده از این پیکره، جهت آشنایی هر چه بیشتر با مفاهیم به کار رفته در آن، مقاله ای را که به عنوان مرجع ارائه شده است حتماً مطالعه فرمایید.**

ابتدا به معرفی ساختار دو کلاسی که جهت نگهداری و بازگرداندن اطلاعات به کار می روند می پردازیم. طبیعتاً با ساختن نمونه از این دو کلاس در برنامه اصلی خود ب راحتی می توانید به تمامی اطلاعات مطلوب خود به تفکیک ویژگی ها دسترسی پیدا نمایید.

۱. کلاس **sentences** جهت نگهداری جملات پیکره با ساختار زیر:

```
public class sentences {
    public sentences() { }
    public string id; // شناسه جمله
    public string text; // متن جمله
    public string polari ty; // بار معنایی جمله
}
```

۲. کلاس با نام **tag** جهت نگهداری اطلاعات برچسب جملات با ساختار زیر:

```
public class tag {
    public tag() { }
    public string id; // شناسه برچسب
    public string type; // نوع برچسب
    public string coordi nate; // مختصات برچسب
    public string relati on; // شناسه پدر
    public string polari ty; // بار معنایی
    public string root; // ریشه
}
```

¹ XML

² DLL (Dynamic Link Library)

توضیحات مربوط به کلاس tag:

- فرمت مربوط به coordinate به صورت زیر است:

[id, beginning character index, ending character index]

- به عنوان مثال اگر برای یک برچسب مقدار coordinate به صورت [gr-1-1, 0, 5] باشد، این بدان معنا است که این برچسب در جمله‌ای با شناسه gr-1-1 قرار دارد و از کاراکتر 0 در جمله آغاز و به کاراکتر 5 در جمله ختم می‌شود.
- مقدار **relation** برای برچسب از نوع نظر^۳ برابر با شناسه پدر آن (یعنی برچسبی از نوع هدف نمونه^۴) است و این مقدار برای برچسبی از نوع هدف نمونه برابر با شناسه پدر آن (یعنی هدف اصلی^۵) است.

توابع موجود در واسط کاربردی

- ✓ در تمامی توابعی که در ادامه آمده‌اند، منظور از **path** رشته‌ای است که آدرس یک سند ایکس.ام.ال را در خود دارد.

```
public List<sentences> getCommentByID(string path, string ID, int choice)
```

این تابع، با دریافت کردن آدرس یک سند ایکس.ام.ال، تمامی جملات یک نظر از نوع نظر متخصص^۶، نظر عمومی^۷ یا یک نظر نقدی^۸ با شناسه مشخص ID را برمی‌گرداند. به عنوان مثال اگر یک نظر با شناسه gr-1 در پیکره شامل سه جمله با شناسه‌های gr-1-1، gr-1-2 و gr-1-3 باشد، با استفاده از این تابع و قرار دادن مقدار پارامتر ID معادل با gr-1، تابع مورد نظر لیستی به ترتیب شامل جملات اول تا سوم این نظر به همراه بار معنایی آنها و همچنین به عنوان مولفه پایانی لیست، کل جملات نظر مربوطه به همراه بار معنایی کلی نسبت داده شده به آن نظر را برمی‌گرداند.

تنها نکته‌ای که در استفاده از این تابع وجود دارد آن است که هرگاه نیاز داشتیم نظر متخصص را به عنوان خروجی دریافت کنیم، مقدار پارامتر ID را برابر با **review** قرار می‌دهیم. یادآوری می‌شود که در هر سند یک نظر که توسط متخصص نوشته شده است موجود است و جملات این بخش معمولاً حالت رسمی‌تری نسبت به دو بخش دیگر دارند.

³ Opinion

⁴ Target(I) or instance target

⁵ Target(M) or main target

⁶ Review

⁷ General review

⁸ Critical review

`public List<sentences> getSenByID(string path, string ID)`

این تابع، با دریافت کردن آدرس یک سند ایکس.ام.ال، یک جمله مشخص از نوع نظر متخصص، نظر عمومی یا یک نظر نقدی با شناسه مشخص ID را برمی‌گرداند.

`public List<sentences> getAllSen(string path, int choice)`

این تابع با دریافت کردن آدرس یک سند ایکس.ام.ال، تمامی جملات موجود در سند مذکور از نوع نظر عمومی یا نظر نقدی به همراه بار معنایی آنها را برمی‌گرداند. جهت انتخاب نوع نظر عمومی مقدار پارامتر choice برابر با 0 و جهت انتخاب نوع نظر نقدی مقدار پارامتر choice برابر با 1 تعیین می‌گردد.

`public List<tag> tagListbyID(string path, string ID)`

این تابع با دریافت کردن آدرس یک سند ایکس.ام.ال و شناسه یک جمله با مقدار ID، تمامی برچسب‌های از نوع نظر یا هدف نمونه متعلق به آن جمله مشخص را برمی‌گرداند.

`public List<tag> keywordListbyID(string path, string ID)`

این تابع با دریافت کردن آدرس یک سند ایکس.ام.ال و شناسه یک جمله با مقدار ID، تمامی برچسب‌های از نوع کلمه کلیدی⁹ متعلق به آن جمله مشخص را برمی‌گرداند.

`public List<tag> targetList(string path)`

این تابع با دریافت کردن آدرس یک سند ایکس.ام.ال، لیستی از تمامی برچسب‌های از نوع هدف اصلی سند مذکور را برمی‌گرداند.

⁹ Keyword

نمونه ای از کد و نحوه استفاده از واسط کاربردی

```
SentiPers.API obj = new API();
List<SentiPers.API.sentences> senList = new List<API.sentences>();
List<SentiPers.API.tag> tags = new List<API.tag>();
string result;
string path = @"D:\corpus\"; //آدرس پوشه قرارگیری اسناد پیکره
DirectoryInfo directory = new DirectoryInfo(path);
FileInfo[] files = directory.GetFiles("*.xml");
foreach (FileInfo file in files)
{
    string name = Path.GetFileNameWithoutExtension(file.Name);
    senList = obj.GetAllSen(path + file.Name, 0);
    result = "";
    tags.Clear();
    for (int i = 0; i < senList.Count; i++)
    {
        result = result + senList[i].id + System.Environment.NewLine;
        result = result + senList[i].text + System.Environment.NewLine;
        result = result + senList[i].polarity + System.Environment.NewLine;
        tags = obj.tagListbyID(path + file.Name, senList[i].id);
        for (int j = 0; j < tags.Count; j++)
        {
            result = result + tags[j].id + "," + tags[j].type + "," + tags[j].coordinate +
            "," + tags[j].relation + System.Environment.NewLine;
        }
    }
    File.WriteAllText(@"D:\test\" + file.Name, result);
}
```

✓ لطفاً در صورت وجود هرگونه مشکل یا پیشنهاد در خصوص استفاده از این واسط کاربردی و توسعه آن با این آدرس

ایمیل در ارتباط باشید: hosseini.p.gu@gmail.com

✓ آدرس وبسایت گروه: <http://nlp.guilan.ac.ir>

گروه پردازش زبان طبیعی دانشگاه گیلان

