

## روش و مراحل اجرای پروژه قافیه یاب با استفاده از اشعار ثبت شده در سایت گنجور

۱- در اولین قدم، دیتابیس گنجور (نسخه ۲/۹۶) از منبع گیت‌هاب آن با لینک زیر دانلود شد

<https://github.com/ganjoor/desktop/releases/tag/v2.96>

۲- تغییرات زیر در دیتابیس اعمال شد

- فیلد kind به جدول poem جهت نوع چینش قافیه اضافه شد
  - i. Kind=1: مثنوی
  - ii. Kind=2: غزل، قصیده، رباعی
  - iii. Kind=3: قطعه
  - iv. Kind=4: ناشناخته (سیستم نتوانسته نوع شعر را تشخیص دهد)
- فیلد ID به جدول verse به صورت AUTOINCREMENT اضافه شد.
- دو فیلد قافیه (rhyming\_words) و ID هم قافیه (rhyme\_group\_id) به جدول Verse اضافه شد.

```
-- اضافه کردن ستون‌ها
ALTER TABLE "verse"
  ADD COLUMN rhyming_words TEXT;

ALTER TABLE "verse"
  ADD COLUMN rhyme_group_id INTEGER;

ایندکس برای سرعت بخشیدن به جستجو و گروه‌بندی بر اساس قافیه
CREATE INDEX idx_verse_rhyme_group
  ON "verse" (poem_id, rhyme_group_id);

CREATE INDEX idx_verse_rhyming_words
  ON "verse" (rhyming_words);
```

۳- با استفاده از یک کد پایتون نوع اشعار kind (چینش قافیه) بر اساس چهار یا شش مصراع اول تشخیص داده شد.

- تشخیص قافیه بر اساس تکرار دو حرف آخر کلمه آخر انجام میشود.
- اگر کلمه (کلمات) آخر دقیقا مثل هم باشد (ردیف داشته باشد) به کلمه قبلی رجوع میکند.
- این روش خطا هم دارد. مثلا «پدر» و «سر» که هم قافیه هستند ولی با توجه به نبود اعراب گذاری، سیستم فقط یک حرف مشابه تشخیص میدهد.
- برای افزایش دامنه تعداد قافیه های تشخیصی، برخی از مواردی که سیستم kind=4 تشخیص داده، به صورت دستی اصلاح شدند. مثلا بر اساس title در جدول poem (مثلا اگر نوشته رباعی، kind=2)

- از تعداد ۱۱۹،۰۶۱ عدد شعر در جدول poem قالب شعر (چینش قافیه) تعداد ۱۰۲،۰۳۰ تشخیص داده شدند.
  - بر این اساس از تعداد ۲،۷۹۰،۸۱۴ مصراع در جدول Verse تعداد ۲،۴۲۵،۸۴۵ مصراع قافیه آنها پیدا شد. و تعداد ۳۶۴،۹۶۹ قافیه یابی نشدند که البته برخی از این مصراع ها اصلا شعر نیستند. (متن یا نثر)
  - در نهایت تعداد ۱۱۵،۶۸۵ کلمه منحصر به فرد قافیه کشف شد.
- ۴- سپس با کد پایتون دیگری، کلمه قافیه هر مصرع تشخیص داده شده و در فیلد rhyming\_words جدول verse ذخیره شد. کلمات هم قافیه یک شعر فیلد rhyme\_group\_id یکسان دارند.
- ممکن است دو کلمه در دو شعر با هم هم قافیه باشند، ولی rhyme\_group\_id یکسانی نداشته باشند. در این روش فقط قافیه هایی که یکبار استفاده شده اند جمع آوری شده و هدف تولید قافیه جدید نیست.
  - لازم است توجه شود که اگر یک کلمه قافیه مشترک بین دو شعر وجود داشته باشد در نتیجه نهایی وبگاه عروض، کلیه هم قافیه های هر دو شعر نمایش داده میشود.
- ۵- پس از آماده سازی دیتابیس و انتقال آن به وبگاه عروض، هم اکنون با جستجو هر کلمه قافیه، هم قافیه های آن و مثال های مختلف برای هر قافیه نمایش داده میشود.

```
SELECT DISTINCT v.rhyming_words, v.text ,p.title, p.url
FROM verse v
JOIN poem p ON v.poem_id = p.id
JOIN verse vv ON v.rhyme_group_id = vv.rhyme_group_id
WHERE vv.rhyming_words = 'بند';
```

خروجی تستی برای یافتن هم قافیه «بند»

لینک به سایت گنجور	عنوان	نمونه شعر	کلمه قافیه
<a href="https://ganjoor.net/ferdousi/shahname/tahmoores/sh1">https://ganjoor.net/ferdousi/shahname/tahmoores/sh1</a>	طهمورث	پسر بد مر او را یکی هوشمند	هوشمند
<a href="https://ganjoor.net/ferdousi/shahname/tahmoores/sh1">https://ganjoor.net/ferdousi/shahname/tahmoores/sh1</a>	طهمورث	گرانمایه طهمورث دیو بند	بند
<a href="https://ganjoor.net/ferdousi/shahname/tahmoores/sh1">https://ganjoor.net/ferdousi/shahname/tahmoores/sh1</a>	طهمورث	هر آن چیز کاندز جهان سودمند	سودمند
<a href="https://ganjoor.net/ferdousi/shahname/tahmoores/sh1">https://ganjoor.net/ferdousi/shahname/tahmoores/sh1</a>	طهمورث	کنم آشکارا گشایم ز بند	بند
<a href="https://ganjoor.net/ferdousi/shahname/tahmoores/sh1">https://ganjoor.net/ferdousi/shahname/tahmoores/sh1</a>	طهمورث	چه آزاد گشتند از بند او	بند
<a href="https://ganjoor.net/ferdousi/shahname/tahmoores/sh1">https://ganjoor.net/ferdousi/shahname/tahmoores/sh1</a>	طهمورث	بمستند ناچار پیوند او	پیوند