



تشخیص خودکار مقام‌های قرآنی با استفاده از یادگیری ماشین

پایان‌نامه کارشناسی ارشد زبان‌شناسی رایانشی

استاد راهنما: دکتر حسین صامتی

استاد راهنمای همکار: دکتر محمد بحرانی

نگارش: محمدجواد خداپنده

انگیزه و کاربردها



- ❖ متمرکز بودن بیشتر پژوهش‌ها
- روی موسیقی
- ❖ بازیابی خودکار محتوا
- ❖ یادگیری الکترونیکی قرآن
- ❖ سامانه‌های توصیه‌گر
- ❖ ...



اهداف



- ❖ تهیه دادگان برچسب خورده برای تشخیص مقام‌های قرآنی
- ❖ بررسی عملکرد ضرایب ام‌اف‌سی‌سی و اچ‌پی‌سی‌پی در بازنمایی اطلاعات صوت
- ❖ سنجش عملکرد روش‌های سنتی یادگیری ماشین شامل ماشین بردار پشتیبان، شبکه عصبی پیش‌خور و شبکه بازگشتی ال‌اس‌تی‌ام در دسته‌بندی مقام‌های قرآنی

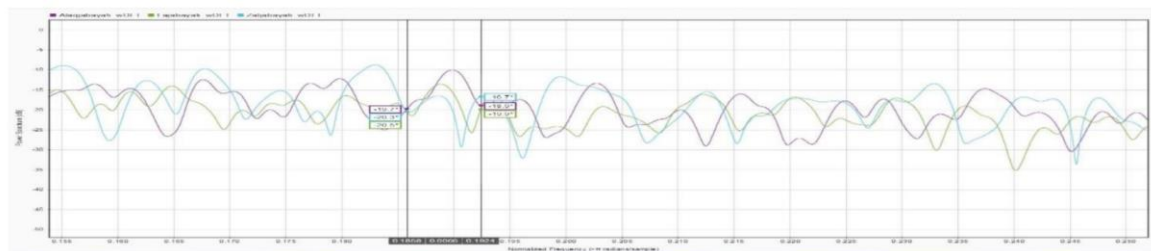
موسیقی تلاوت



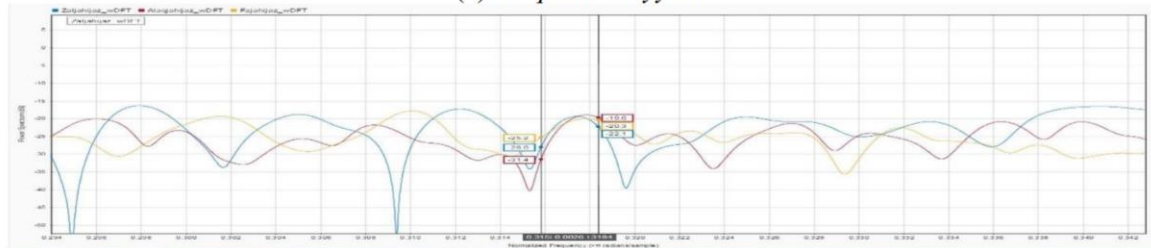
- ❖ اکتاو: فاصله دو نت با نسبت بسامد ۲ یا ۵.
- ❖ گام: نت‌هایی که به صورت متصل (متوالی) میان فاصله اکتاو هستند.
- ❖ جنس: برش‌های ۳، ۴ یا ۵ نتی از یک گام با فواصل ثابت
- ❖ مقام: از اتصال اجناس به دست می‌آیند و تقسیم‌بندی‌های مختلفی دارند.
- ❖ انتقال: پرده‌گردانی از یک جنس به جنس دیگر به واسطه نت غماز
- ❖ تصویر: بالا بردن یا پایین آوردن تمام نت‌های یک جنس به میزان چند درجه
- ❖ ...



تشخیص مقام‌های قرآنی



(a) Maqamat Bayyati



مقام	سیکاه	صبا	رست	حجاز	بیاتی	عجم	جمع
آموزش	۳۴۶	۲۹۰	۶۸۳	۱۶۸	۹۵۶	۱۲۸	۲۵۷۱
آزمایش	۸۷	۷۳	۱۷۱	۴۲	۲۳۹	۳۲	۶۴۴
جمع	۴۳۳	۳۶۳	۸۵۴	۲۱۰	۱۱۹۵	۱۶۰	۳۲۱۵

- ویژگی‌ها: ۱۳ ضریب ام‌اف‌سی‌سی به‌همراه ویژگی‌های دلتا و دلتا-دلتا ام‌اف‌سی‌سی
- بهترین دقت: ۶۲.۷۳ درصد
- بهترین مقادیر برای پارامترهای اس‌وی‌ام (هسته RBF):
 $C = 10; \gamma = 0.001$

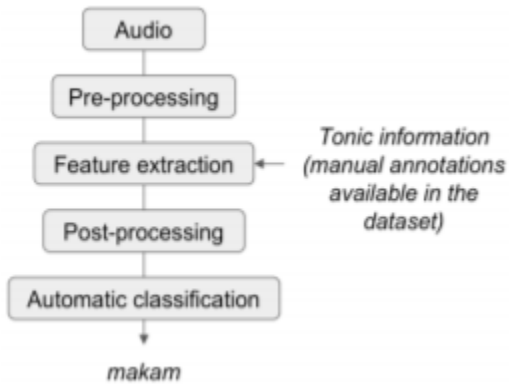
- داده: ۲۱ قطعه تلاوت ضبط‌شده به‌صورت استودیویی (هر سه تلاوت مربوط به یکی از ۷ خانواده مقامی)
- هدف: پروفایل‌بندی مقامات قرآنی و مقایسه اطلاعات به دست آمده از طیف توان سیگنال‌های مربوط به هر خانواده مقامی

Jabar, F. H., Mohammad, J. I., Faizal M. Zain, A., Hasan, A. B., Jalil, D. M., Mohd, R. A., . . . Ahmad, H. (2018). PRELIMINARY RESULTS OF QURANIC MAQAMAT PROFILING USING W-DFT TECHNIQUE. 25th International Congress on Sound and Vibration: Hiroshima Calling (pp. 4572-4579). Hiroshima, Japan: International Institute of Acoustics and Vibration (IIAV).



خدابنده، م. و بحرانی، م. (۱۳۹۸). تشخیص خودکار مقام‌های قرآنی با استفاده از ماشین‌های بردار پشتیبان. اولین همایش ملی هوش مصنوعی و محاسبات نرم در علوم انسانی. تهران.

تشخیص مقام‌های موسیقی ترکی



“RBF kernel”
 $C = \{0.001, 0.01, 0.1, 1, 10, 100, 1000\};$
 $\gamma = \{0.001, 0.01, 0.1, 1\}$

Makam Type	#_of Tracks	Makam Type	#_of Tracks	Makam Type	#_of Tracks
Acemaşiran	50	Huzzam	50	Rast	50
Acemkürdi	49	Karcigar	50	Saba	50
Bestenigar	50	Kurdilihicazkar	50	Segah	50
Beyati	49	Mahur	50	Sultaniyegah	50
Hicaz	50	Muhayyer	50	Suzinak	50
Hicazkar	50	Neva	50	Ussak	50
Huseyni	49	Nihavent	50	Total	997

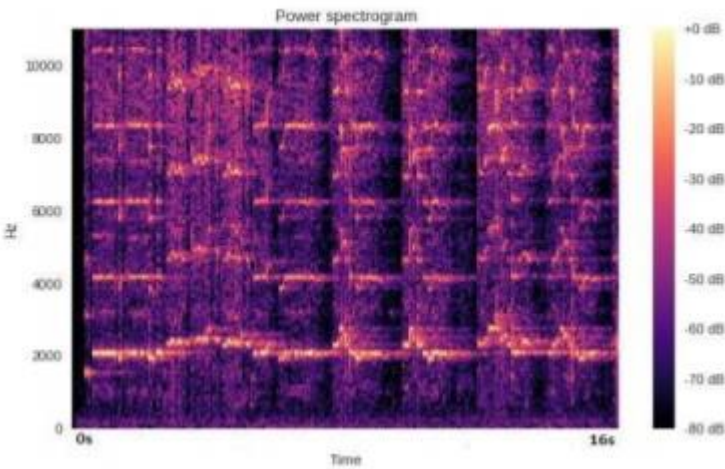
■ ویژگی‌ها: میانگین + انحراف معیار بردارهای
 اچ‌پی‌سی‌پی کل و قسمتی از قطعات در طول زمان
 (۲۰۰ میلی‌ثانیه با ۱۰۰ میلی‌ثانیه هم‌پوشانی، بازه
 فرکانس ۱۰۰ تا ۵۰۰۰ هرتز)

نتایج →

<u>F-Measures</u>	12 - bins	24 - bins	36 - bins	48 - bins
Mean	0.64	0.64	0.65	0.66
Stdev.	0.65	0.7	0.69	0.7
Mean+Std	0.65	0.7	0.7	0.7



تشخیص گوشه‌ها و دستگاه‌های موسیقی ایرانی



Name of Dastgah	Precision	Recall	F1
Shour	97.42	87.52	92.21
Homayoun	76.22	82.72	79.34
Mahour	91.53	89.92	90.72
Segah	83.58	84.95	84.26
Chahargah	63.04	91.53	74.66
Rastpanjgah	90.66	90.30	90.48
Nava	96.12	87.92	91.84

- **دادگان: ۱۱۳۷** قطعه صوتی ۱۶ ثانیه‌ای
(فرکانس نمونه‌برداری: ۸۱۹۲ هرتز)
نواخته‌شده با ویولن و نی (سازهای اصلی)
- **ویژگی‌ها:** طیف به‌دست‌آمده از تبدیل فوریه
زمان سریع



دادگان پیشنهادی

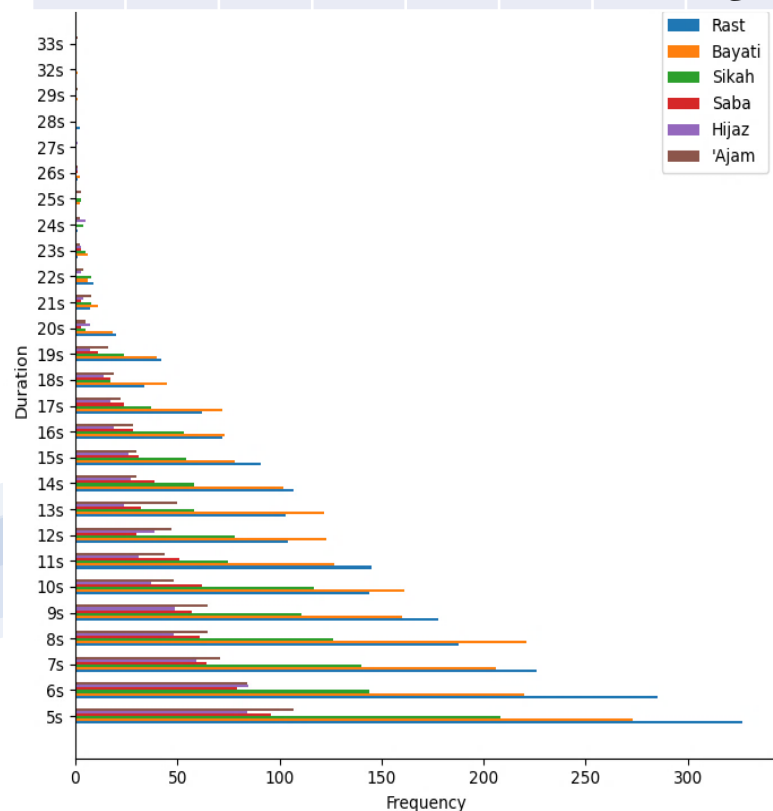


جمع	سیکاه	صبا	رست	حجاز	بیاتی	عجم	
۴۲۶۰	۸۰۴	۴۱۰	۱۳۸۱	۲۰۵	۱۳۱۹	۱۴۱	منشای
۳۳۲۸	۵۲۹	۲۸۲	۷۶۸	۳۸۵	۷۵۱	۶۱۳	شحات
۷۵۸۸	۱۳۳۳	۶۹۲	۲۱۴۹	۵۹۰	۲۰۷۰	۷۵۴	منشای + شحات
۷۲۹	۱۳۴	۷۱	۲۳۴	۳۵	۲۳۰	۲۵	منشای
۵۹۶	۹۶	۵۰	۱۳۴	۶۹	۱۳۶	۱۱۱	شحات
۱۳۲۵	۲۳۰	۱۲۱	۳۶۸	۱۰۴	۳۶۶	۱۳۶	منشای + شحات

حاشیه‌نویسی و قطعه‌بندی فایل‌ها با نرم‌افزار آداسیتی
هر قطعه، در یک مقام خاص و در یک نفس اجرا شده
است.

معیار برچسب‌زنی: وجود جنس اصلی هر خانواده
مقامی در قطعه

قطعات با طول زیر ۵ ثانیه حذف شدند.



آماده‌سازی داده

انتخاب قاری

حاشیه‌نویسی

بارگیری صوت

پیش‌پردازش و استخراج ویژگی

تغییر فرمت به دلیوای وی

تغییر نرخ نمونه‌برداری به ۱۶ کیلوهرتز

تبدیل به صوت تک‌کاناله

حذف سکوت ابتدا و انتها

زیرنمونه‌گیری

استخراج ضرایب
ام‌اف‌سی سی

استخراج بردارهای
اچ‌پی سی پی

میانگین‌گیری
از ویژگی‌ها
در طول زمان

(ex., seg., feat.)

(fr., feat.)

(ex., feat.)

تنظیم ابرپارامترهای شبکه بازگشتی ال‌اس‌تی‌ام

تنظیم ابرپارامترها

مدل اولیه شبکه
بازگشتی ال‌اس‌تی‌ام

تنظیم ابرپارامترهای شبکه پیش‌خور

تنظیم ابرپارامترها

مدل اولیه شبکه
پیش‌خور

پارامترهای نهایی ویژگی‌ها

پارامترها و مدل نهایی
اس‌وی‌ام

تنظیم ابرپارامترهای
استخراج ویژگی و اس‌وی‌ام

جستجوی شبکه‌ای
ابرپارامترها

مدل اولیه اس‌وی‌ام

مدل نهایی شبکه پیش‌خور

مدل نهایی شبکه بازگشتی
ال‌اس‌تی‌ام

آزمایش مدل‌ها

تنظیم توأم پارامترهای استخراج ویژگی و اسویام

مقادیر کاندیدا

پارامترهای ام‌اف‌سی‌سی	پارامترهای اچ‌بی‌سی‌بی	طول پنجره (هزارم‌ثانیه): {۳۰، ۲۰}
		مقدار هم‌پوشانی (هزارم‌ثانیه): {صفر، ۱۰}
		بیشینه بسامد فیلتربانک (هرتز): {۸۰۰۰، ۵۰۰۰}
		وضوح اکتاو (بین): {۴۸، ۱۵۹}
پارامترهای اسویام		مقدار «گاما» در هسته پایه‌شعاعی ^۱ : {۰.۱، ۰.۰۱، ۰.۰۰۵، ۰.۰۰۱}
		ضریب هموارسازی اسویام ^۲ : {۱، ۵، ۱۰، ۲۰، ۳۱، ۶۰، ۱۰۰، ۱۰۰۰}

■ تعیین بهترین مقادیر برای پارامترها با جستجوی شبکه‌ای

■ برای هر حالت، اعتبارسنجی متقابل ۵-فولد لایه‌ای با ۲ تکرار استفاده شد.

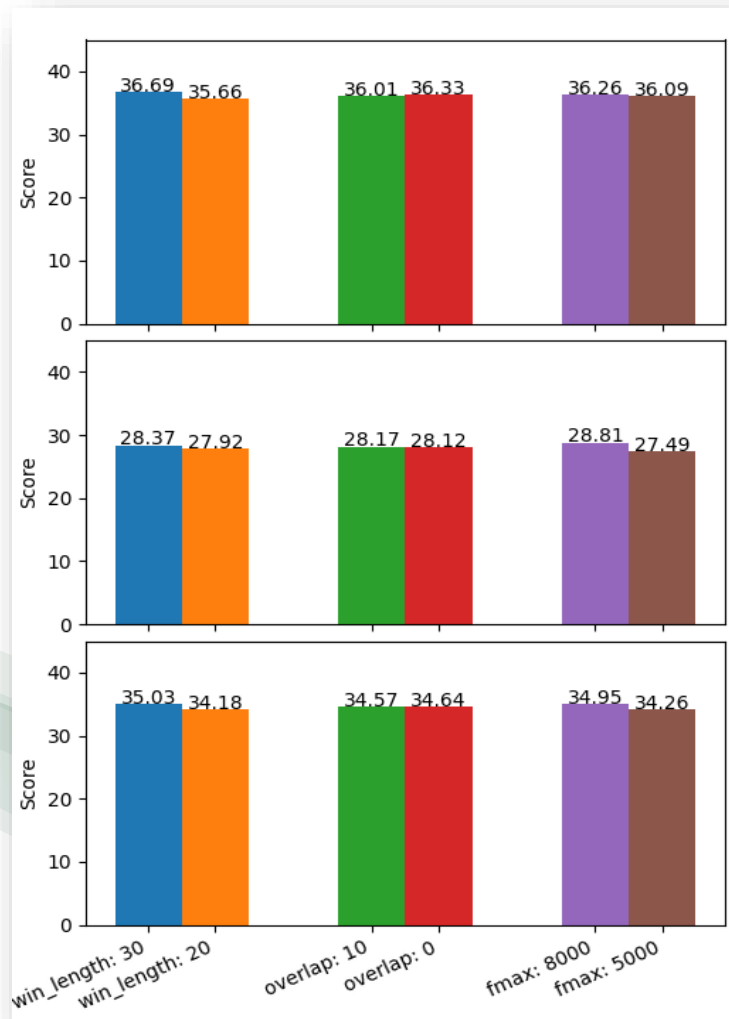


تنظیم پارامترهای استخراج ویژگی

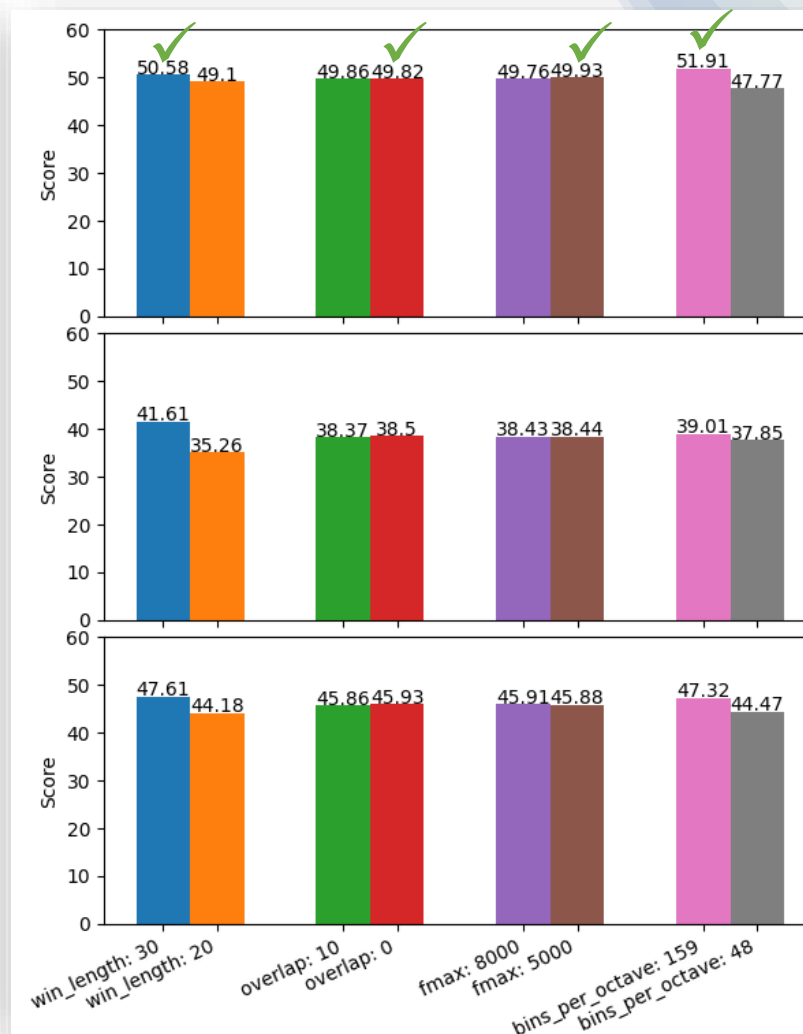
منشأ

شحات

کل دادگان



امتیاز مقادیر مختلف برای پارامترهای امافسی



امتیاز مقادیر مختلف برای پارامترهای کروما



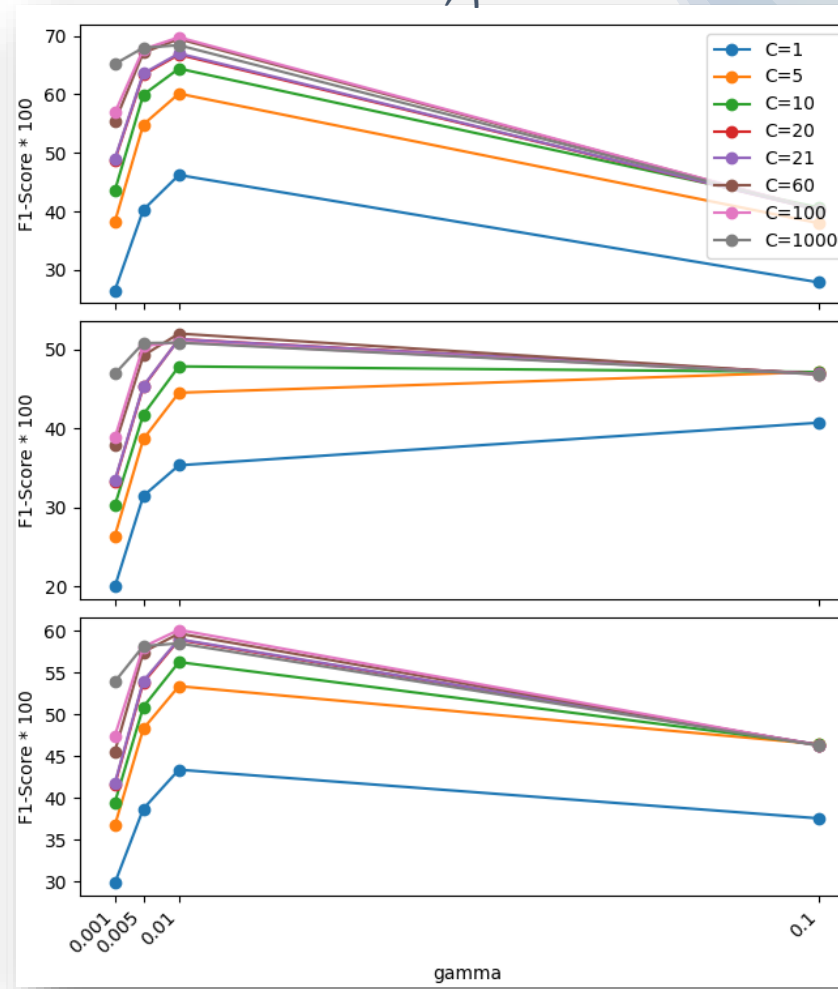
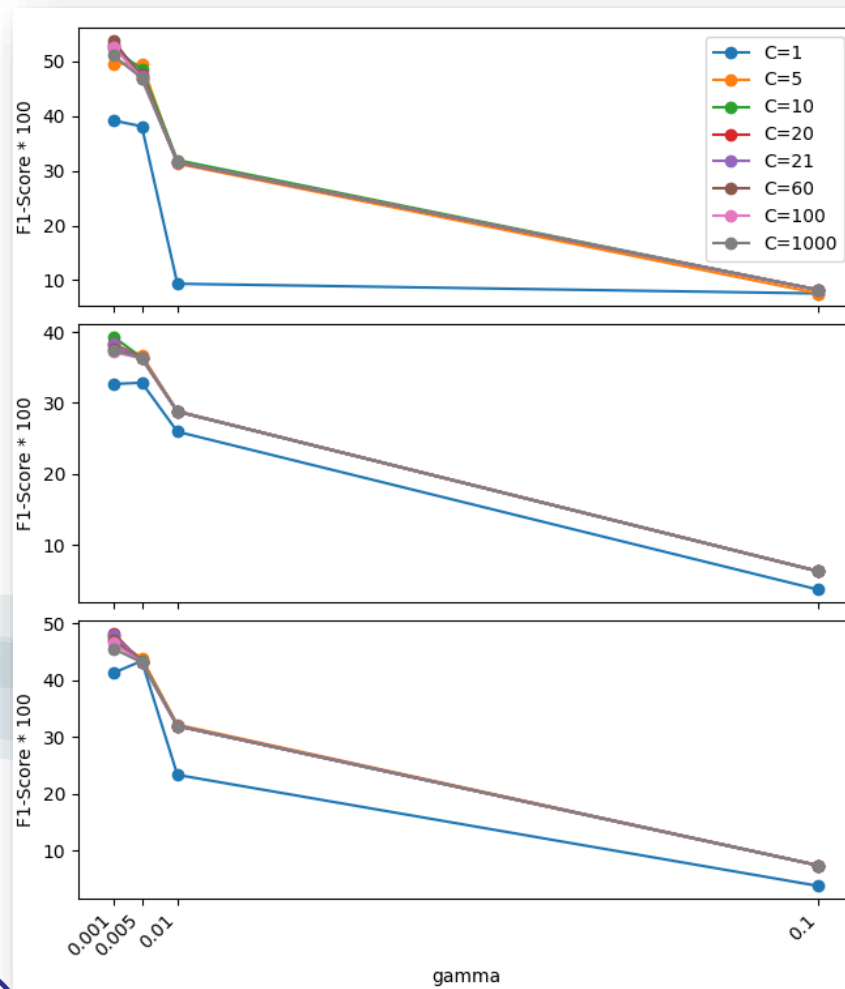
تنظیم پارامترهای اس وی ام

✓ $C = 100$; $\gamma = 0.01$ ✓

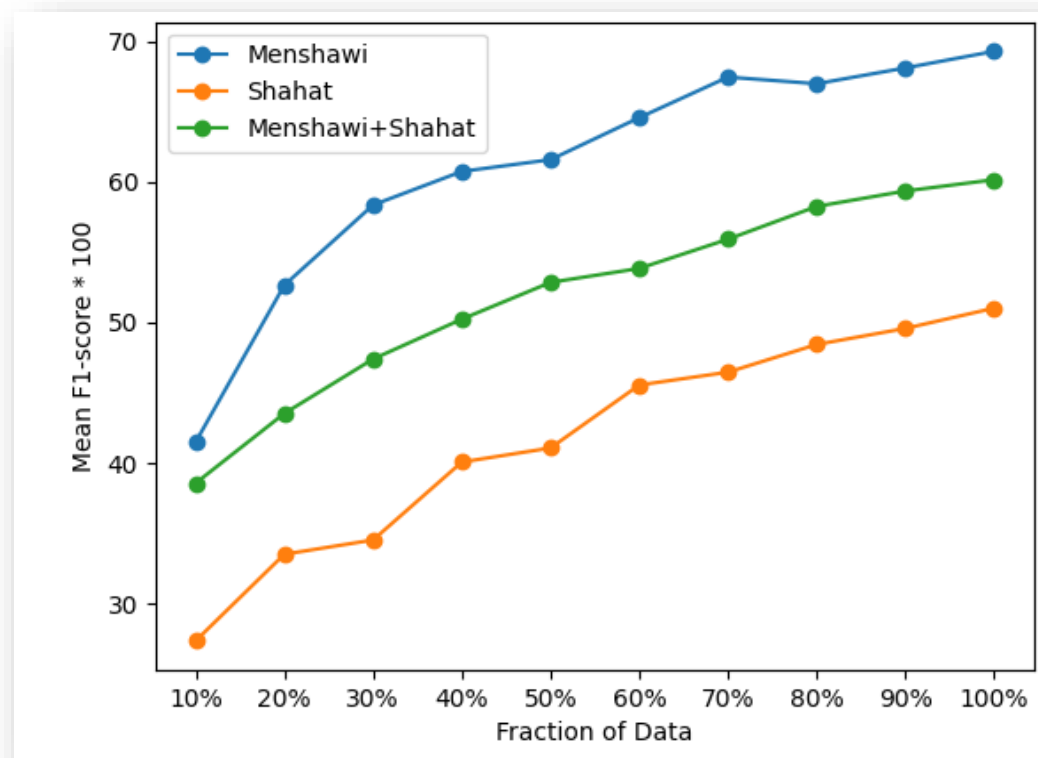
منشأوی

شحات

کل دادگان



تأثير افزایش داده‌ها



خلاصه نتایج – اس وی ام

تلاوت منشاوی			تلاوت شحات		مجموع دادگان	
دسته‌ها	امتیاز-اف دسته‌بند	دسته‌ها	دسته‌ها	امتیاز-اف دسته‌بند	دسته‌ها	امتیاز-اف دسته‌بند
عجم و صبا	۹۶.۲۷	حجاز و راست	دسته‌ها	۸۵.۹۵	حجاز و راست	۸۷.۸۸
عجم، حجاز و سیکاه	۸۷.۸۳	عجم، حجاز و صبا	دسته‌ها	۷۴.۶۴	بیاتی، صبا و سیکاه	۸۰.۰۹
عجم، حجاز، صبا و سیکاه	۸۲.۹۷	بیاتی، حجاز، صبا و سیکاه	دسته‌ها	۶۵.۱۱	بیاتی، حجاز، راست و صبا	۷۲.۶۳
عجم، بیاتی، حجاز، صبا و سیکاه	۷۸.۹۷	عجم، بیاتی، حجاز، صبا و سیکاه	دسته‌ها	۵۵.۳۴	عجم، بیاتی، حجاز، صبا و سیکاه	۶۶.۶
عجم، بیاتی، حجاز، راست، صبا و سیکاه	۷۲.۷۱	عجم، بیاتی، حجاز، راست، صبا و سیکاه	دسته‌ها	۵۱.۷۳	عجم، بیاتی، حجاز، راست، صبا و سیکاه	۵۹.۹۶



آماده سازی داده

انتخاب قاری

بارگیری صوت

حاشیه نویسی

پیش پردازش و استخراج ویژگی

تغییر فرمت به دبیو ای وی

تغییر نرخ نمونه برداری به ۱۶ کیلوهرتز

تبدیل به صوت تک کاناله

حذف سکوت ابتدا و انتها

زیر نمونه گیری

استخراج ضرایب
ام اف سی سی

استخراج بردارهای
اچ پی سی پی

میانگین گیری
از ویژگی ها
در طول زمان

پارامترهای نهایی ویژگی ها

پارامترها و مدل نهایی
اس وی ام

تنظیم ابرپارامترهای شبکه پیش خور

مدل اولیه شبکه
پیش خور

تنظیم ابرپارامترها

مدل نهایی شبکه پیش خور

تنظیم ابرپارامترهای شبکه بازگشتی ال اس تی ام

مدل اولیه شبکه
بازگشتی ال اس تی ام

تنظیم ابرپارامترها

مدل نهایی شبکه بازگشتی
ال اس تی ام

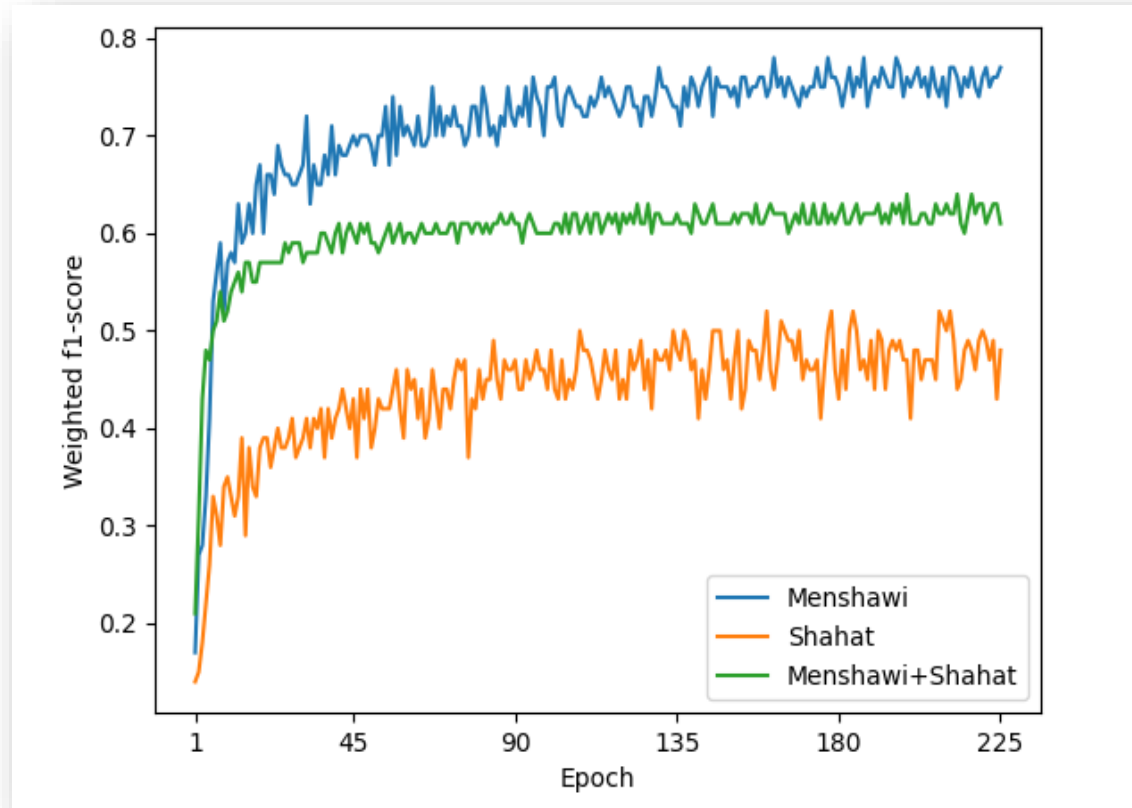
آزمایش مدل ها

تنظیم ابرپارامترهای استخراج ویژگی و اس وی ام

جستجوی شبکه ای
ابرپارامترها

مدل اولیه اس وی ام

خلاصه نتایج — شبکه پیش خور



دسته‌بند شبکه پیش خور	مجموع دادگان			تلاوت شحات			تلاوت منشاوی		
	امتیاز-اف	فراخوانی	صحت	امتیاز-اف	فراخوانی	صحت	امتیاز-اف	فراخوانی	صحت
	۶۱.۵۷	۵۹.۷	۶۸.۱۳	۴۹.۸۳	۴۹.۸	۵۳.۱	۷۵.۱۱	۷۵.۱۲	۷۶.۲



آماده سازی داده

انتخاب قاری

حاشیه نویسی

بارگیری صوت

پیش پردازش و استخراج ویژگی

تغییر فرمت به دلیوای وی

تغییر نرخ نمونه برداری به ۱۶ کیلوهرتز

تبدیل به صوت تک کاناله

حذف سکوت ابتدا و انتها

زیر نمونه گیری

استخراج ضرایب
ام اف سی سی

استخراج بردارهای
اچ پی سی پی

میانگین گیری
از ویژگی ها
در طول زمان

(ex., seg., feat.)

(fr., feat.)

(ex., feat.)

تنظیم ابر پارامترهای شبکه بازگشتی ال استی ام

مدل اولیه شبکه
بازگشتی ال استی ام

تنظیم ابر پارامترها

تنظیم ابر پارامترهای شبکه پیش خور

مدل اولیه شبکه
پیش خور

تنظیم ابر پارامترها

آزمایش مدل ها

مدل نهایی شبکه بازگشتی
ال استی ام

مدل نهایی شبکه پیش خور

پارامترها و مدل نهایی
اس وی ام

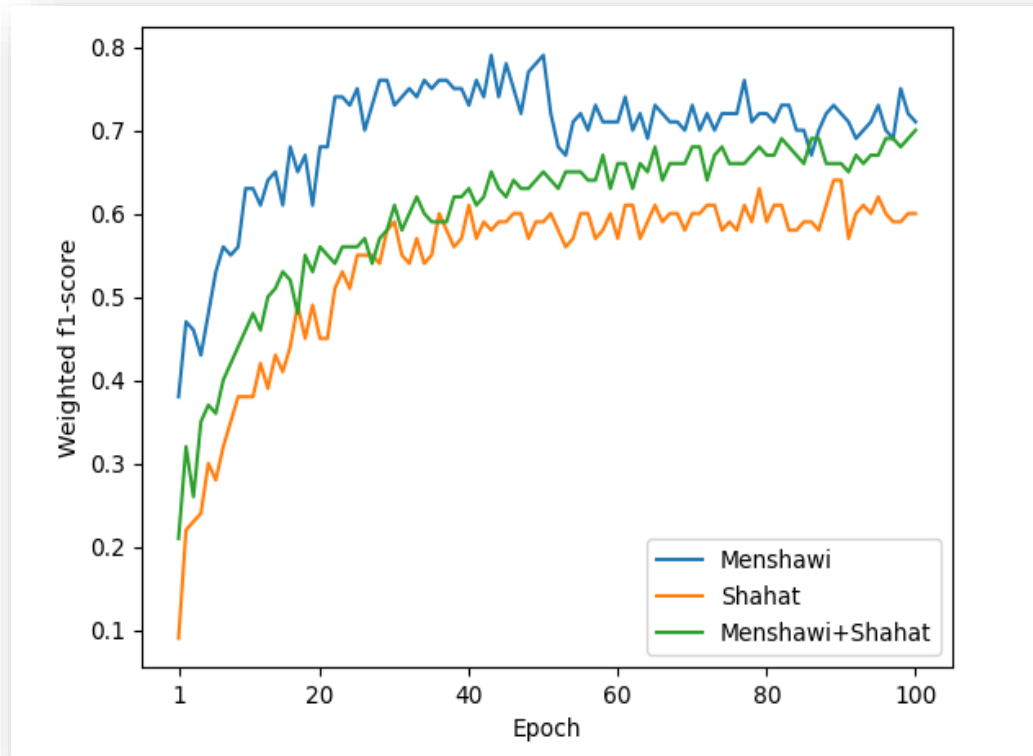
پارامترهای نهایی ویژگی ها

تنظیم ابر پارامترهای
استخراج ویژگی و اس وی ام

جستجوی شبکه ای
ابر پارامترها

مدل اولیه اس وی ام

خلاصه نتایج — شبکه ال اس تی ام



دسته‌بند شبکه ال اس تی ام	مجموع دادگان			تلاوت شحات			تلاوت منشاوی		
	امتیاز-اف	فراخوانی	صحت	امتیاز-اف	فراخوانی	صحت	امتیاز-اف	فراخوانی	صحت
	۶۹.۷۵	۶۹.۳۶	۷۲.۰۸	۶۱.۲۲	۶۱.۶	۶۳.۰۲	۸۱.۴۸	۸۰.۷۵	۸۱.۴۸



- تهیه دادگان برچسب خورده با کیفیت از چند قاری مختلف
- استفاده از روش‌های **self-supervised** در صورت کمبود دادگان برچسب خورده
- استخراج ویژگی‌های بیشتر
- هم‌طول کردن قطعات با استفاده از روش‌هایی مانند DTW
- طراحی سامانه برخط تشخیص مقام
- تشخیص مقام‌های ترکیبی، اصلی و فرعی



[۱] خدابنده، م. و بحرانی، م. (۱۳۹۸). تشخیص خودکار مقام‌های قرآنی با استفاده از ماشین‌های بردار پشتیبان. اولین همایش ملی هوش مصنوعی و محاسبات نرم در علوم انسانی. تهران.

[2] Jabar, F. H., Mohammad, J. I., Faizal M. Zain, A., Hasan, A. B., Jalil, D. M., Mohd, R. A., . . . Ahmad, H. (2018). PRELIMINARY RESULTS OF QURANIC MAQAMAT PROFILING USING W-DFT TECHNIQUE. 25th International Congress on Sound and Vibration: Hiroshima Calling (pp. 4572-4579). Hiroshima, Japan: International Institute of Acoustics and Vibration (IIAV).

[3] Demirel, E., Bozkurt, B., & Serra, X. (2018). Automatic Makam recognition using chroma features. 8th International Workshop on Folk Music Analysis, (pp. 19-24). Thessaloniki.

[4] Rezazadeh Azar, S., Ahmadi, A., Malekzadeh, S., & Samami, M. (2018, 12 17). Instrument-Independent Dastgah Recognition of Iranian Classical Music Using AzarNet. arXiv. Retrieved from <https://arxiv.org/abs/1812.07017>



با تشکر از توجهتان!