



۱. (۱۵٪) [پژوهش: تجزیه معنایی و ترکیب‌پذیری] هدف این تمرین آشنایی با تجزیه معنایی^۱ است. در تجزیه معنایی هدف، دریافت

معنا و منطق جمله‌های زبان طبیعی است، برخلاف تجزیه نحوی که ساختار یک جمله را بر اساس قواعد دستور زبان بررسی می‌کند و در رویارویی با جملات مبهم، ناقص، یا پیچیده به مشکل می‌خورد، تجزیه معنایی یک جمله را به یک نمایش منطقی تبدیل می‌کند که معنا را دریافت و منطق جمله را نشان می‌دهد و می‌توان از آن برای کاربردهای مختلف، مانند ترجمه ماشینی، پرسش و پاسخ^۲، تولید کد و غیره استفاده کرد. همچنین تجزیه معنایی ترکیب‌پذیر^۳ می‌تواند جمله‌های پیچیده و تودرتو را مدیریت کند اما نیازمند حجم زیادی از داده‌گان برچسب‌خورده، دانش خاص حوزه^۴ و به مدل‌هایی نیاز دارد که از عهده ترکیب‌پذیری^۵ و تعمیم‌پذیری^۶ برآیند.

- به توانایی ساخت معانی پیچیده از معانی ساده‌تر با به کارگیری قواعد ترکیب، ترکیب‌پذیری گفته می‌شود، به عنوان مثال، در تجزیه معنایی، ترکیب‌پذیری به این معناست که معنای یک جمله را می‌توان از معانی اجزای آن و نحوه ترکیب آن‌ها استخراج کرد.
- به توانایی به کارگیری دانش یا مهارت‌های آموخته شده در موقعیت‌های جدید یا دیده نشده، تعمیم‌پذیری می‌گویند، به عنوان مثال، در تجزیه معنایی، تعمیم‌پذیری به این معناست که یک تجزیه‌گر معنایی بدون دیدن داده‌گان در زمان آموزش بتواند آن‌ها را به بازنمایی منطقی صحیح متناظر نگاشت کند.

شعر شکلی از زبان طبیعی است که اغلب از هنجارهای دستور زبان منحرف می‌شود، از زبان تمثیلی^۷ استفاده می‌کند و به صدا و ریتم بستگی دارد. هرچند که این ویژگی‌ها شعر را جذاب و پرمعنا می‌کنند اما نتیجه دشوار شدن تجزیه معنایی است. تجزیه معنایی شعر می‌تواند به ما در کشف معانی ضمنی، احساسات و زیبایی‌شناسی اشعار کمک کند. اخیراً، مدل‌های زبانی بزرگ از پیش آموزش دیده، در وظیفه تجزیه معنایی عملکرد قابل توجهی از خود نشان داده‌اند.

۱. یک مقاله اخیر (از سال ۲۰۲۰ تا کنون) در مورد این موضوع مطالعه کرده و بیان کنید که چگونه از مدل‌های زبانی بزرگ برای

تولید تجزیه معنایی اشعار استفاده می‌شود. همچنین خلاصه‌ای از ایده اصلی^۸، روش‌شناسی^۹، انگیزه^{۱۰} و نتایج آن ارائه کرده و در نهایت نقاط قوت و ضعف آن را از دید خودتان ارزیابی کنید (نیازی به تسلط به مفاهیم و جزئیات نظری روش مانند شبکه‌های عصبی برای این سوال ندارید).

۲. (۲۰٪) [نظری: تشخیص احساس و مدل مخفی مارکوف] هدف از این بخش تمرین بر مدل مخفی مارکوف و استفاده از این روش به

شکل مشابه در برچسب‌زنی اجزای کلام برای حل دیگر وظایف در پردازش زبان طبیعی است. قصد داریم با یک مجموعه از برچسب‌های احساس، هر کلمه در پیکره زیر را برچسب‌گذاری کنیم.

« نان بوی خوش‌مزهای می‌دهد، » « این قهوه بوی تلخ و افتضاح داشت »

مجموعه برچسب‌های $\{+, -, 0\}$ در اختیار ما قرار داده شده که به ترتیب بیانگر احساس مثبت، منفی و خنثی است. با در نظر گرفتن فرض مارکوف و استقلال خروجی (به این معنا که هر برچسب فقط به برچسب قبلی و هر کلمه فقط به برچسب متناظر خود وابسته است) یک مدل مارکوف پنهان روی برچسب‌ها و کلمه‌ها خواهیم داشت.

پارامترهای زیر از مدل مخفی مارکوف در اختیار ما قرار گرفته است:

¹ Semantic Parsing

² Question Answering

³ Compositional Semantic Parsing

⁴ Domain Specific Knowledge

⁵ Compositionality

⁶ Generalization

⁷ Figurative Speech

⁸ Objective

⁹ Methodology

¹⁰ Motivation

آشنایی با زبان‌شناسی رایانشی (۱۲۱-۰۵-۸۳)

نیم‌سال دوم ۱۴۰۱-۱۴۰۲



تاریخ تحویل: ۱۴۰۲/۰۳/۲۶

تمرین شماره ۴

دانشکده علوم و فنون نوین

و	افتضاح	تلخ	نان	قهوه	خوشمزه	بو	این	داشت	می‌دهد	
۰.۰	۰.۰	۰.۱	۰.۱	۰.۱	۰.۶	۰.۱	۰.۰	۰.۰	۰.۰	+
۰.۰	۰.۵	۰.۳	۰.۱	۰.۰	۰.۰	۰.۲	۰.۰	۰.۰	۰.۰	-
۰.۰۵	۰.۰۵	۰.۰۵	۰.۱	۰.۱	۰.۰۵	۰.۱	۰.۱۵	۰.۱۵	۰.۱۵	0

جدول ۱: احتمالات توزیع، ϕ که به طور مشخص در $\phi(y_j|x_j)$ سطرها برچسب‌های پنهان (z_j) و ستون‌ها مشاهدات (x_j) را نشان می‌دهند.

0	-	+	
۰.۷	۰.۲	۰.۱	∅
۰.۴	۰.۳	۰.۳	+
۰.۳	۰.۵	۰.۲	-
۰.۶	۰.۲	۰.۲	0

جدول ۲: احتمالات انتقال، θ ، سطرها برچسب‌های z_j و ستون‌ها z_{j+1} هستند. انتقال از حالت تهی نشان‌دهنده احتمال اولیه حالت است.

۱. ساختار مدل مخفی مارکوف فوق را رسم کنید.
۲. در صورتی که دنباله‌ی برچسب‌ها «0، 0، 0، +» (از راست به چپ خوانده شود) باشد و عبارت «نان بوی خوش‌مزه‌ای دارد» را داشته باشیم، احتمال توام این رخداد را حساب کنید.
۳. در صورتی که دنباله‌ی «قهوه بوی تلخ می‌دهد» به ما داده شده باشد، به کمک الگوریتم ویتربی برچسب‌های احساس مناسب این دنباله را حساب کنید.

۳. (۶۵٪) [پایاده‌سازی: تجزیه‌گر نحوی احتمالی] هدف از این بخش پیاده‌سازی یک تجزیه‌گر PCFG^۱ است که از الگوریتم ویتربی برای یافتن محتمل‌ترین درخت تجزیه‌ی برای یک جمله و یک گرامر PCFG استفاده می‌کند. مجموعه دادگان بانک درخت وابستگی جهانی فارسی در اختیار شما قرار داده شده است که بر اساس بانک درخت وابستگی فارسی ساخته شده، در الگوی طراحی بانک درخت وابستگی برچسب‌ها ریز-دانه^۲ و برای زبان خاصی است اما بانک درخت وابستگی جهانی از الگوی جهانی^۳ و برچسب‌زنی بین‌زبانی^۴ پیروی می‌کند. این تمرین در ۴ بخش: بخش ۱: آماده‌سازی دادگان، بخش ۲: تخمین PCFG، بخش ۳: تجزیه CKY با رمزگشایی ویتربی و بخش ۴: ارزیابی آماده شده است.

۳-۱ آماده‌سازی دادگان

هدف از این بخش تبدیل داده‌های خام به یک قالب مناسب برای تجزیه‌گر شما است. در این بخش، شما دادگان [PerUDT](#) را از وبسایت [وابستگی‌های جهانی](#) دانلود کرده و برای تجزیه‌گر خود پیش‌پردازش می‌کنید. دادگان شامل سه فایل `fa_perdt-ud-train.conllu`، `fa_perdt-ud-dev.conllu` و `fa_perdt-ud-test.conllu` که به ترتیب مجموعه‌های آموزش، توسعه و آزمون را شامل می‌شوند، هر فایل در قالب CoNLL-U^۵ است که یک قالب جداشده با تب است که هر توکن را در یک خط و هر جمله را در یک پاراگراف نشان می‌دهد.

¹ Probabilistic Context Free Grammar

² Fine-grained

³ Universal

⁴ Cross-linguistic

⁵ [CONLL-U Format](#)



۱. تابعی بنویسید که یک فایل CoNLL-U را می‌خواند و یک لیست از جمله‌ها را بر می‌گرداند، که هر جمله یک لیست از دوتایی‌ها به شکل (کلمه و برچسب) است، به عنوان مثال؛ جمله «این کتاب را بخوانید.» در فایل آموزش باید به صورت زیر نشان داده شود:

```
[ ('این', 'DET'), ('کتاب', 'NOUN'), ('را', 'ADP'), ('بخوانید', 'VERB'), ('.', 'PUNCT') ]
```

۲. در ادامه لازم است تابعی بنویسید که یک لیست از جمله‌ها را به لیستی از درخت‌ها تبدیل کند، که هر درخت به صورت یک لیست تودرتو از غیر-پایانه‌ها و پایانه‌ها نشان داده شده است. به عنوان مثال؛ جمله «این کتاب را بخوانید.» در فایل آموزش باید به صورت زیر تبدیل شود:

```
[ 'S', [ 'NP', [ 'این', 'DET'], [ 'کتاب', 'NOUN'] ], [ 'POSTP', [ 'را', 'ADP'] ], [ 'VP', [ 'بخوانید', 'VERB'] ], [ '.', 'PUNCT'] ]
```

شما می‌توانید از هر کتابخانه موجود یا ماژول مانند NLTK، anytree و غیره برای انجام این تبدیل استفاده کنید.

۳-۲ تخمین PCFG

هدف از این بخش یادگیری گرامر احتمالی از درخت‌های برچسب‌زده شده است. تابعی بنویسید که از لیستی از درخت‌ها، یک گرامر PCFG تخمین می‌زند. گرامر PCFG شامل یک مجموعه از قواعد به شکل $X \rightarrow YZ$ یا $X \rightarrow w$ که در آن X, Y, Z و w غیر-پایانه‌ها و پایانه است، که در آن هر قاعده دارای احتمالی‌ست که مشخص می‌کند که چقدر احتمال دارد برای بازنویسی X استفاده شود، در نظر داشته باشید که مجموع احتمال تمام قواعد با LHS یکسان باید به یک برسد.

۱. تابعی بنویسید که لیستی از درخت‌ها را به عنوان ورودی بگیرد و نگاشتی از قاعده-احتمال برگرداند. شما باید از تخمین حداکثر درست‌نمایی^۲ برای محاسبه احتمالات استفاده کنید، به این معنا که تعداد رخداد هر قاعده در درخت‌ها را بشمارید و بر تعداد دفعات رخ دادن LHS آن تقسیم کنید. به عنوان مثال، اگر شما ۱۰۰ درخت دارید و قاعده $S \rightarrow NP VP$ به تعداد ۵۰ بار و S ۸۰ بار رخ داده باشد، پس احتمال $S \rightarrow NP VP$ برابر با ۰.۶۲۵ است. به صورت ریاضی، می‌توانید به صورت زیر بنویسید:

$$q(X \rightarrow \alpha) = \frac{\text{Count}(X \rightarrow \alpha)}{\text{Count}(X)}$$

که در آن

$$\text{Count}(X \rightarrow \alpha)$$

تعداد دفعات رخداد قاعده $X \rightarrow \alpha$ در درخت‌ها و $\text{Count}(X)$ تعداد دفعات رخداد X به عنوان LHS در درخت‌ها است. این معادله به این معناست که برای پیدا کردن احتمال یک قاعده، ما باید فرکانس آن را با فرکانس LHS آن مقایسه کنیم. به عنوان مثال، اگر ما ۱۰۰ درخت داشته باشیم و مشاهده کنیم که ۵۰ بار $S \rightarrow NP VP$ و ۳۰ بار S رخ داده است، پس می‌توانیم محاسبه کنیم:

$$q(S \rightarrow NP VP) = \frac{\text{Count}(S \rightarrow NP VP)}{\text{Count}(S)} = \frac{50}{50 + 30} = 0.625$$

$$q(S \rightarrow NP) = \frac{\text{Count}(S \rightarrow NP)}{\text{Count}(S)} = \frac{30}{50 + 30} = 0.375$$

¹ Left-Hand Side

² Maximum Likelihood Estimation



توجه کنید که همانطور که باید، مجموع این احتمالات به یک می‌رسد.

۲. همچنین لازم است تابعی دیگری بنویسد که نگاشتی از قواعد-احتمالات را به عنوان ورودی می‌گیرد و آن‌ها را در قالب خوانایی

نشان می‌دهد. به عنوان مثال، شما می‌توانید هر قاعده را در خط جداگانه با احتمال آن در پرانتز نشان دهید، برای مثال:

S -> NP VP (0.625)

S -> NP (0.375)

NP -> DET NOUN (0.8)

NP -> NOUN (0.2)

....

۳-۳ تجزیه CKY با رمزگشایی ویتربی

هدف از این بخش یافتن بهترین درخت تجزیه برای یک جمله داده شده با استفاده از دستور زبان شما است. در این بخش، تابعی لازم است بنویسد که الگوریتم CKY را با رمزگشایی ویتربی پیاده‌سازی می‌کند تا برای جمله و دستور زبان PCFG داده شده، محتمل‌ترین درخت تجزیه را پیدا کند. الگوریتم CKY یک الگوریتم با رویکرد برنامه‌نویسی پویا است که یک جدول مثلثی را با بهترین مولفه‌ها برای هر بخش از کلمات در جمله پر می‌کند. رمزگشایی ویتربی گسترشی از الگوریتم CKY است که نشان‌گرهای پشتیبان^۱ را برای هر سلول در جدول نگه می‌دارد و برای بازسازی بهترین درخت تجزیه در پایان استفاده می‌شود.

۱. تابعی بنویسد که یک جمله (لیست کلمه‌ها) و یک گرامر (نگاشتی از قواعد-احتمالات) را به عنوان ورودی می‌گیرد و بهترین درخت تجزیه (لیست تودرتو از غیر-پایانه‌ها و پایانه‌ها) را به عنوان خروجی بر می‌گرداند (از لگاریتم برای جلوگیری از کمبود حساسی آ هنگام ضرب احتمالات استفاده کنید. شما همچنین باید کلمات ناشناخته را با اختصاص یک برچسب دلخواه مانند X مدیریت کنید، همچنین برای پیاده‌سازی الگوریتم تجزیه CKY و روش Viterbi به تابع‌های دیگری نیاز دارید، که متناسب با طراحی خود باید آن‌ها را نیاز پیاده‌سازی کنید).

۳-۴ ارزیابی

هدف از این بخش سنجش عملکرد تجزیه‌گر شما بر روی داده‌های دیده نشده است. شما تجزیه‌گر خود را بر روی مجموعه آزمون با استفاده از معیارهای استاندارد مانند دقت^۲، بازخوانی^۳، صحت^۴ و امتیاز^۵ F1^۶ ارزیابی خواهید کرد. دقت به عنوان نسبت جملاتی که تجزیه‌گر شما درخت تجزیه صحیح را (مانند PerUDT) برمی‌گرداند، تعریف می‌شود. بازخوانی به عنوان نسبت مؤلفه‌های صحیح (غیر-پایانه‌ها یا پایانه‌ها) در PerUDT که توسط تجزیه‌گر شما نیز یافت می‌شوند، تعریف می‌شود. صحت به عنوان نسبت مؤلفه‌های یافت شده توسط تجزیه‌گر شما که در PerUDT نیز صحیح هستند، تعریف می‌شود. امتیاز F1 به عنوان میانگین هارمونیک بازخوانی و صحت تعریف می‌شود.

۱. تابعی بنویسد که دو لیست از درخت‌ها (یکی از PerUDT و یکی از تجزیه‌گر شما) را به عنوان ورودی می‌گیرد و چهار عدد (دقت، بازخوانی، صحت و امتیاز F1) را به عنوان خروجی برمی‌گرداند، لازم است برای اطمینان از خروجی نتایج خود، این نتایج را به کمک کتابخانه‌های یا ماژول‌های آماده نیز ارائه کنید.

¹ Backpointers

² Underflow

³ Accuracy

⁴ Recall

⁵ Precision

⁶ F-1 Score



۲. همچنین این معیارها را برای تجزیه‌گر خود بر روی هر دو مجموعه توسعه و آزمون گزارش کنید و آن‌ها را با برخی نتایج پایه مقایسه کنید مانند همیشه بازگرداندن یک درخت خالی یا همیشه بازگرداندن یک درخت دلخواه.

دلیل این که شما نیاز به استفاده از مجموعه توسعه دارید، تنظیم هایپرپارامترهای^۱ تجزیه‌گر شما و جلوگیری از بیش‌برازش^۲ به مجموعه آموزش است. مجموعه توسعه زیرمجموعه‌ای از داده‌ها است که برای آموزش استفاده نمی‌شود بلکه برای اعتبارسنجی و بهینه‌سازی است. شما می‌توانید از مجموعه توسعه برای مقایسه نسخه‌های مختلف تجزیه‌گر خود و انتخاب بهترین آن بر اساس عملکرد آن استفاده کنید. با این حال، شما نباید از مجموعه توسعه برای ارزیابی تجزیه‌گر نهایی خود استفاده کنید، زیرا ممکن است تحت تاثیر انتخاب‌هایی که در حین تنظیم کردید، مغرض باشد. به جای آن، شما باید از مجموعه آزمون استفاده کنید، که زیرمجموعه‌ای دیگر از داده‌ها است که تا آخر ننگه داشته شده و فقط یک بار برای سنجش قدرت تعمیم‌پذیری تجزیه‌گر شما استفاده می‌شود.

در ادامه مواردی آورده شده است که از آن‌ها می‌توانید در طراحی تجزیه‌گر بهره ببرید:

- از [نمونه آنلاین هضم](#) می‌توانید برای بررسی پردازش، تحلیل صرفی و تجزیه نحوی خود استفاده کنید.
- به کمک [ابزار آنلاین](#) می‌توانید تجزیه وابستگی جمله‌های بانک درختی که در اختیار شما قرار داده شده است را مشاهده کنید.
- به کمک ابزار [displaCy](#) می‌توانید بصری‌سازی تجزیه نحوی جمله‌های خود را دریافت کنید.
- به کمک ابزار [syntree](#)، می‌توانید درخت تجزیه نحوی خود را نمایش دهید.
- برای مطالعه بیشتر در مورد روش‌های تجزیه می‌توانید کدهای [کتابخانه جعبه ابزار زبان طبیعی](#)^۳ را بررسی کنید.

۴. (نمره اضافه ۳۰٪) [پایاده‌سازی: سنتز گفتار و شناسایی فرمت‌ها] هدف از این بخش آشنایی با سیستم‌های ساخت گفتار، بررسی طیف‌نگاره^۴ و فرمت‌ها است.

قصه داریم که تعدادی عبارت را با صدای خود و با صدای گویندگان سرویس‌های ساخت گفتار، سنتز کنیم. برای ضبط صدای خود می‌توانید از ابزارهای مختلفی استفاده کنید، از ابزارهای رایج می‌توان آدیشن^۵، اوداسیتی^۶ و ... را نام برد. همچنین سرویس ناراکیٹ^۷ از جمله سرویس‌های سنتز گفتاری است که از گویندگان فارسی بهره می‌برد.

۱. لازم است که جمله‌واره‌های زیر را با صدای خود و با کمک ابزارهای سنتز صدا جداگانه ضبط کنید. خروجی خواسته‌شده با کمک سرویس‌های سنتز صدا را حتما با دو گوینده مرد و زن به صورت جدا سنتز شود. همچنین نام فایل‌های سنتز شده حتما باید مشخص، متمایز و با قاعده باشند، به عنوان مثال؛ در سرویس ناراکیٹ گوینده مرد (که در آن سرویس با «رضا» شناخته می‌شود) با کد 01 و گوینده زن (که در آن سرویس با «شهره» شناخته می‌شود) با کد 02 از هم متمایز می‌شوند، که در نهایت صدای سنتز شده می‌تواند با NARAKEET_SENTENCE_SPEAKER نام‌گذاری شود (توجه: در صورتی که از دیگر سرویس‌های سنتز صدا به جز ناراکیٹ استفاده می‌کنید، حتما باید توضیح کامل از تفاوت آن با ناراکیٹ، دلیل استفاده و روش استفاده از آن را در گزارش خود ذکر کنید).

¹ Hyperparameters

² Overfit

³ NLTK

⁴ Spectrogram

⁵ Adobe Audition

⁶ Audacity

⁷ [Narakeet](#)

آشنایی با زبان‌شناسی رایانشی (۱۲۱-۰۵-۸۳)
نیم‌سال دوم ۱۴۰۱-۱۴۰۲



تاریخ تحویل: ۱۴۰۲/۰۳/۲۶

تمرین شماره ۴

شماره جمله	جمله
۱	دانشگاه تهران
۲	کارخانه ساخت کلاه‌خود، کلاه خود و خودرو
۳	گوینده اوستا
۴	هر گردی گردو و گردویی گرد

۲. خطاها یا کاستی‌هایی که در سنتزها رخ داده است را در گزارش خود بیان کنید.
۳. برای هر جمله‌واره به‌صورت تک‌واژ با خطاها برخورد کنید و خطای نهایی را برای هر جمله‌واره به صورت تجمعی با معیار «نرمال شده خطا در تک‌واژ» (به بیان دیگر؛ مجموع همه خطاها در همه تک‌واژهای آن جمله‌واره، تقسیم بر شمارگان همه تک‌واژهای آن جمله‌واره) گزارش کنید.
۴. به کمک تبدیل فوریه زمان-کوتاه^۱ طیف‌نگاره عبارت‌ها را که برای سنتزهای مختلف آماده شده، بسازید (توجه؛ می‌توانید برای ساخت جمله‌واره‌ها از تکنیک آوانگاری فونوتیک^۲ بهره ببرید). شبه‌کد^۳ زیر جهت انجام تبدیل فوریه زمان-کوتاه در اختیار شما قرار گرفته است.

```
# Set the parameters for the STFT
window_size = 1024
hop_size = 512
fft_size = 1024

# Create a window function to reduce spectral leakage
window = create a hanning window of size window_size

# Create an empty list for the spectrogram
spectrogram = []

# For each frame of the audio data with a hop size
for each frame in data with a step of hop_size:
    # Apply the window function to the frame
    frame = multiply frame and window element-wise

    # Compute the spectrum of the frame using FFT
    spectrum = do FFT on frame with size fft_size

    # Keep only the positive frequencies of the spectrum
    spectrum = get the first half of spectrum

    # Get the magnitude of the spectrum
    magnitude = get the absolute value of spectrum

    # Add the magnitude to the spectrogram list
    add magnitude to spectrogram

# Convert the spectrogram list to a matrix
spectrogram = convert spectrogram to a matrix

# Show the spectrogram as an image
show spectrogram as an image
```

۵. برای همه‌ی صدای‌های سنتز شده فرمنت اول را مشخص کرده و مقدار آن را در گزارش ارائه کنید، F_0 صدای شما چقدر است؟

¹ Short-Time Fourier Transform
² Phonotic Transcription
³ Pseudocode

آشنایی با زبان‌شناسی رایانشی (۱۲۱-۰۵-۸۳)
نیم‌سال دوم ۱۴۰۱-۱۴۰۲



تاریخ تحویل: ۱۴۰۲/۰۳/۲۶

تمرین شماره ۴

دانشکده علوم و فنون نوین

۶. برای هر طیف‌نگاره محل رخداد هر واج را مشخص و با طیف‌نگاره‌های متناظر (سنتز شده با سرویس سنتز صدا مانند نراکیت) مقایسه کنید.

در هنگام سنتز صدا، موارد زیر را در نظر بگیرید:

- فرکانس نمونه‌برداری، اندازه چندی‌سازی و تعداد کانال را ۱۶ کیلوهرتز، ۱۶-بیت و تک‌کاناله باشد.
- فایل صداهای ضبط شده را هم طول کنید (در زمان). برای این منظور، پیشنهاد می‌شود تا به فایل با مدت زمان کمتر، در آغاز و پایان آن، به مقدار برابر سکوت اضافه شود (برای مثال: اگر فایل مدنظر یک ۱ ثانیه کوتاه‌تر است، ۰.۵ ثانیه به آغاز و ۰.۵ ثانیه به پایان اضافه شود).



راهنمای تحویل

- ❖ انجام تمرین به صورت تک نفره است.
- ❖ جهت تحویل تمرین می‌بایست پیاده‌سازی‌ها از پایه و بدون استفاده از کتابخانه‌های موجود انجام شود، اما در شرایط قید کتابخانه در صورت سوال، منعی ندارد.
- ❖ زمان مناسبی را برای نوشتن و آماده‌سازی گزارش در نظر بگیرید، چرا که نیمی از نمره هر سوال، بر اساس گزارش تحویلی است.
- ❖ در صورت استفاده از **کدهای آماده**، دلیل استفاده، کامنت‌گذاری و توضیحات کافی، ضروری است، در غیر اینصورت تقلب تلقی می‌گردد.
- ❖ گزارش‌های تمرین خود را در مسیر `documentation` و اطلاعات مرتبط با پیاده‌سازی‌ها نیز در مسیر `source` قرار داده شوند.
- ❖ لطفاً گزارش، فایل‌کدها و سایر ضمیمه‌ها را با فرمت `CL_YourFamilyName_YourStNo_HW#.zip` به ایمیل h.veisi@ut.ac.ir ارسال فرمائید.
- ❖ برای اطلاعات بیشتر به [صفحه درس](#) به آدرس <https://dsp.ut.ac.ir/courses/y1401/introduction-to-computational-linguistics> مراجعه کنید.

در صورت وجود سوال، ابهام و درخواست راهنمایی در گروه اسکایپی یا تلگرامی و یا از طریق ایمیل با دستیار آموزشی در ارتباط باشید.