**EROL ÖZKAN**

**Konuşma Tanıma**

**-Sphinx-**

**İÇİNDEKİLER**

# 1.Training Languange Model

# 2.Building Phonelist

# 2.Building Dictionary

# 4.Training Acustic Model

# 5. Testing Using Pocketsphinx

# 6. Testing Using Java

# 7. Results

**1 – Training Languange Model :**

* İlk aşamada TextCorpus.txt isimli dosyasındaki tüm satırların başlarına ve sonlarına; basit bir java kodu yazılarak “<s>” ve “</s>” harf dizileri eklendi.
* *Terminal’ de “text2wfreq < TextCorpus.txt > turkish.tmp.vocab”* komutu çalıştırılarak vocabulary dosyası oluşturuldu.
* *“text2idngram -vocab turkish.vocab -idngram turkish.idngram < turkish.closed.txt”*

*“ idngram2lm -vocab\_type 0 -idngram turkish.idngram -vocab turkish.vocab -arpa turkish.lm”* komutu çalıştırılarak arpa formatında languange model oluşturuldu.

* Bir önceki aşamadan gelen languange model dosyası *“Sphinx\_lm\_convert -i turkish.lm -o turkish.lm.bin”* komutu çalıştırılarak binary formatına çevrildi.

**2- Building PhoneList :**

Tüm kelimelerde geçen her harfi içeren bir text dosyası hazırlandı. Türkçe karakterler ii oo uu ss gibi harflerle ifade edildi.

**2 – Building Dictionary :**

Hem TextCorpus.txt’daki tüm kelimeler hem de ses dosyalarının açıklamasında bulunan tüm kelimeler; basit java kodları yazılarak tek bir vocab dosyasında toplandı. Daha sonra bu kelimelerin yanlarına okunuşları ( k e l ii m e şeklinde) yine başka bir java kodu yazılarak eklendi.

**4 – Training Acustic Model :**

* Terminalde *“sphinxtrain -t an4 setup”* komutu çalıştırılarak sphinx’in gerekli dosyaları oluşturması sağlandı.
* Klasorde gerekli dosyalar oluşturuldu.

$CFG\_DICTIONARY = "$CFG\_LIST\_DIR/$CFG\_DB\_NAME.dic";

$CFG\_RAWPHONEFILE = "$CFG\_LIST\_DIR/$CFG\_DB\_NAME.phone";

$CFG\_FILLERDICT = "$CFG\_LIST\_DIR/$CFG\_DB\_NAME.filler";

$CFG\_LISTOFFILES = "$CFG\_LIST\_DIR/${CFG\_DB\_NAME}\_train.fileids";

$CFG\_TRANSCRIPTFILE = "$CFG\_LIST\_DIR/${CFG\_DB\_NAME}\_train.transcription"

* “Sphinxtrain run” komutu çalıştırılarak train aşaması yapıldı.
* En son sphinxtrain -s decode run komutu çalıştırıldı.
* Train sette kullanılacak olan .wav formatindaki ses dosyaları “wav” klasörüne atıldı. Bu ses dosyalarının açıklamaları ise basit bir java kodu yazılarak tek bir text dosyasında toplandı. Hazırlanan bu dosyadaki tüm satırların başlarına “<s>”, sonlarına ise “</s>” eklendi.
* Wav dosyalarının ilk deneme aşamasında sadece 1 tanesi test için bırakılırken, 2. Deneme aşamasında 30 adedi test için bırakıldı. Test set için kullanılan bir ses dosyası train sette yer almadı.

**5. Testing Using Pocket Sphinx**

Training aşamalarında elde edilen çıktılar; *“pocketsphinx\_continuous.exe -inmic yes -lm turkish.lm -dict turkish.dic -hmm turkish.cd\_cont\_200”* yazılarak direk olarak linux üzerinde ses tanıma performansı denendi. (continuous olarak)

**6. Testing Using Java**

Elde edilen çıktılar java platorformu üzerinde sphinx\_4 ile denendi. IDE olarak [IntelliJ IDEA](https://www.jetbrains.com/idea/) kullanıldı. Maven üzerinden dependency’ler eklendi ve hem contious olarak çalışan hem de input olrak file alıp, onu transcribe eden bir java programı yazıldı.

Configuration parametreleri aşağıdaki gibi ayarlandı;

configuration.setAcousticModelPath("OUTPUT/turkish.cd\_cont\_untied");

configuration.setDictionaryPath("OUTPUT/turkish.dic");

configuration.setLanguageModelPath("OUTPUT/turkish.lm");

**7. Results**

1. deneme aşamasında sadece 1 adet ses dosyası test için bırakılırken, 2. Deneme aşamada 30 adet test için bırakıldı. Test set için kullanılan bir ses dosyası train sette yer almadı.

1. Deneme sonucu : 1 test dosyası

SENTENCE ERROR: 100.0% (1/1) WORD ERROR RATE: 60.0% (3/5)

1. Deneme sonucu : 30 test dosyası

SENTENCE ERROR: 96.8% (30/31) WORD ERROR RATE: 512.0% (855/167)