

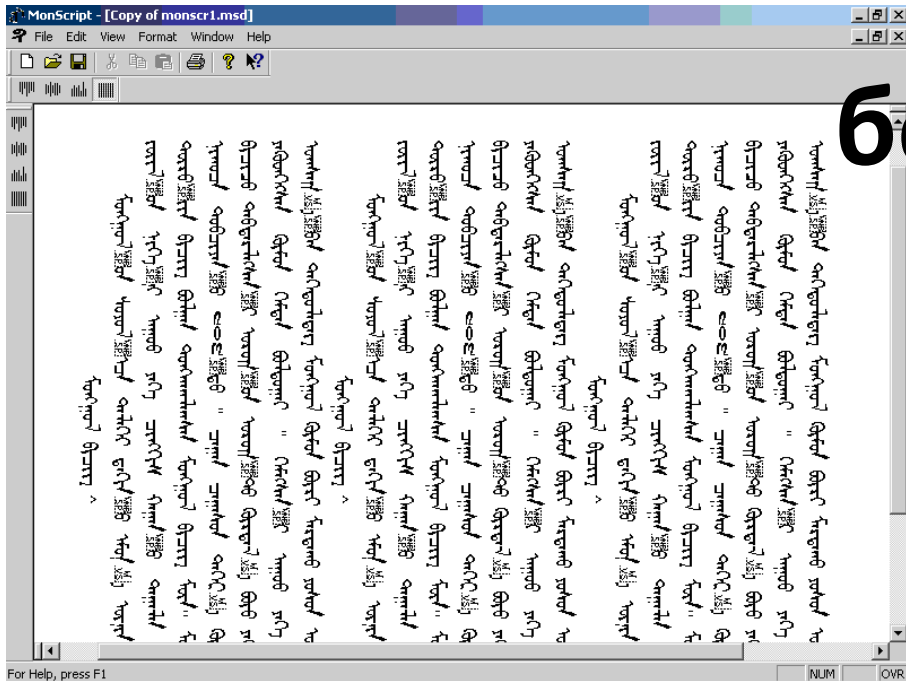
# Текст

## боловсруулалтын үндэс

Лекц №5

Өгүүлбэр сегментлэл  
ба Шийдвэрийн мод

2020 он



# Өгүүлбэр сегментлэл

- !, ? нь харьцангуй салаа утгагүй
- Цэг “.” нь салаа утгатай
  - Өгүүлбэр дуусгана
  - Товчлолд Inc. эсвэл Dr.
  - Тоонд .02% эсвэл 4.3
- 2-тын ангилагч
  - цэг “.” хайна
  - EndOfSentence/NotEndOfSentence аль болохыг шийднэ
  - Ангилагч: гар аргаар бичсэн дүрэм, регуляр илэрхийлэл, эсвэл машин сургалт

# Хэрэв үг нь end-of-sentence гэдгийг тодорхойлно: Шийдвэрийн мод



# Шийдвэрийн модны илүү нарийн онцлог шинж

- Цэгтэй үг нь: Эхний үсэг том, жижиг, бүгд том үсэг, тоо
- Цэгийн дараагийн үг: Эхний үсэг том, жижиг, бүгд том үсэг, тоо
- Тоон шинжүүд
  - Цэгтэй үгийн урт
  - Магадлал(цэгтэй үг өгүүлбэрийн төгсгөлд байх)
  - Магадлал(цэгийн ард өгүүлбэрийн эхэнд үг байх)

# Шийдвэрийн модыг хэрэгжүүлэх нь

- Шийдвэрийн мод нь энгийн if-then-else илэрхийлэл
- Сонирхолтой судалгаа нь онцлог шинжүүдийг сонгох
- Ерөнхийдөө шийдвэрийн модны бүтцийг гар аргаар бий болгоход үргэлж хэцүү байдаггүй.
  - Ингэвэл зөвхөн энгийн шинж чанрууд болон энгийн хүрээнд ашиглана
    - Тоон шинжүүд хувьд, хязгаарыг зааж өгөхөд их хэцүү байдаг.
  - Үүний оронд, бүтэц нь ихэвчлэн сургалтын корпусаас машин сургалтаар суралцдаг

# Шийдвэрийн мод ба бусад ангилагчид

- Дурын ангилагчид ашиглагдаж болох онцлог шинжүүдийг тодорхойлно
  - Логик регресс
  - Тулгуур векторын машин - SVM
  - Нейроны Сүлжээ
  - гэх мэт.