

A gépi fordítás és a neurális gépi fordítás

Machine and neural machine translation

In this globalized world, machine translation has become a daily tool in our lives. In the last few years, 'neural network' has become one of the most widespread expressions in a variety of academic papers. Methods applying the neural network paradigm have brought about scientific breakthroughs in several fields, including computational linguistics and, in particular, machine translation. The notion of the neural network was developed in the 1940s, but the performance of computers at the time was not up to the task of processing the vast amount of data neural networks required. Today's high-performance computers can take advantage of the benefits of neural networks. In this paper the main techniques of machine translation are described, with special emphasis on neural machine translation.

Keywords: machine translation, neural machine translation, statistical machine translation, rule-based machine translation, example-based machine translation

Bevezetés

A gépi fordítás használata elterjedt a hétköznapokban. Az elmúlt évek során a gépi fordító rendszerek jelentős változásokon mentek keresztül. A gépi fordítás a számítástechnika tudományával egyidős. Kezdetben az emberi fordítást hivatott kiváltani, de hamar rájöttek a tudósok, hogy ez nem ilyen egyszerű feladat. Nehézsége abban rejlik, hogy míg egy ember több évtizednyi tudással, tapasztalattal, kreativitással, asszociációs képességgel és egyéb tulajdonságokkal rendelkezik, addig egy gép csak azt fogja tudni, amire megtanítjuk. Komoly kihívás a betanítás, amely számtalan problémával jár. Ilyen például a szavak jelentésbeli, illetve szerkezeti többértelműsége, a különböző nyelvtani szerkezetekkel járó problémák, a szórendek, igeidők stb. Manapság egyre több magánszemély és vállalat használja a gépi fordító szoftvereket. Mind a magánszemélyek, mind a cégek számára nagy segítséget nyújt egy jó minőségű gépi fordító rendszer. Azonban számtalanszor tapasztaljuk, hogy a gépi fordító gyenge minőségű fordítást állít elő. Bár a kezdeti cél az volt, hogy az emberi munkát kiváltsa, jelenleg a gépi fordítás legfőbb célja a megértés, az információkinyerés, valamint az emberi munka támogatása. Ezért amikor gépi fordítás minőségéről beszélünk, az emberi fordítással eltérően, nem stilisztikai hibákra gondolunk, hanem általános nyelvhelyességi hibákról.

A gépi fordítás

A gépi fordítás története (Hutchins–Somers 1992) a 17. században kezdődött, amikor René Descartes egy univerzális nyelv megalkotásáról gondolkodott, amelynek segítségével az egész világon minden ember megértené egymást. Descartes univerzális nyelvének elmélete szótáron, egyértelmű jelentésű kifejezéseken és logikai elveken alapszik. Ez az elmélet hasonlít

a később használt interlingua (lásd „A szabályalapú gépi fordítás” fejezetben) elméletére. Maga a gépi fordítás tudománya a számítógépek megjelenésével egyidős (Hutchin–Somers 1992) és mind a mai napig fontos kutatási terület a számítógépes nyelvészetben. A gépi fordítás fejlődése nagy utat tett meg. Kezdetben, az 1930-as években nyelvtani szabályokkal és szótárakkal közelítették a problémát. Majd később, az 1980-as években megszületett a gondolat, hogy a már létező fordításokat fel lehet használni, így jött létre két új módszer: a példaalapú és a statisztikai gépi fordítás. Az évek során, a gépi fordítás minőségének javítására hibrid módszerekkel is kísérleteztek, különböző gépi fordító módszerek előnyeit kombinálva hoztak létre rendszereket.

Az elmúlt években a gépi fordítás jelentős mértékben fejlődött és számos szoftveres módszer született a természetes nyelvek közötti fordítás megoldására. Mind a három (szabályalapú, példaalapú és statisztikai) fő módszer a mai napig fejlődik és jelentős eredményeket érnek el benne, azonban kétség kívül, a legjelentősebb áttöréseket a 2014-ben megjelent új módszer, a neurális gépi fordítás (Cho–Merri–Bahdanau–Bengio 2014a) érte el.

A jelen tanulmányban a gépi fordítás főbb módszereit mutatom be, majd külön kitérek a neurális gépi fordítás működési elvére.

A szabályalapú gépi fordítás

A szabályalapú gépi fordító (*Rule Based Machine Translation*) (Koehn 2010) beépített szótár és nyelvtani (főként szintaktikai és morfológiai) szabályok alapján végez közvetlen fordítást. Legegyszerűbb formája a direkt fordítás (*direct translation*), ahol a gépi fordító rendszer egy szótár segítségével szóról szóra fordítja le a forrásnyelvi szöveget a célnyelvre, majd a végén, a célnyelvi oldalon, szórendi átalakításokkal javítja az eredményt. A módszer előnye, hogy könnyen megvalósítható, azonban hátránya, hogy nem képes komplex nyelvtani szerkezeteket kezelni, ezért sok esetben rossz minőségű fordítást állít elő. A direkt fordítás módszerénél jobb fordítást eredményez, ha a szótár és az átrendezési szabályok mellé nyelvtani szabályokat is mellékelünk. A fordítás történhet transzfer (közvetlenül a két nyelv között) és interlingua (közvetítő nyelven keresztül) módszerekkel. A forrásnyelvi szövegből különböző komplexitású szabályok segítségével egy köztes reprezentációt hozunk létre (ez az elemzés művelete), majd a köztes reprezentációból újabb előre definiált szabályok segítségével állítjuk elő (ez a generálás művelete) a célnyelvi fordítást. Minél mélyebb az elemzés és a generálás szintje, annál jobb minőségű fordítást tudunk előállítani. A mélység alatt azt értjük, hogy minél többféle elemzést végzünk magán a szövegen. Az elemzés lehet morfológiai, szintaktikai vagy szemantikai elemzés. Ezt az elméletet szemlélteti a Vauquois-háromszög (lásd 1. ábra).

A szabályalapú gépi fordító rendszerhez szükségünk van különböző nyelvi elemző rendszerekre, ami lehet morfológiai, szintaktikai vagy szemantikai elemző. Minél jobb minőségű elemzők állnak a rendelkezésünkre, annál pontosabb fordítást tudunk készíteni. A szabályalapú módszereknek azonban az egyik legnagyobb hátránya, hogy a fordítás minősége erősen függ az elemzők minőségétől, és mivel az egyes típusú elemzések önmagukban is kutatási területek, sajnos kevés jó minőségű elemző áll a rendelkezésünkre. A másik probléma, hogy ezek az elemzők nyelvspecifikusak, ezért megnehezítik a rendszer számára az újabb nyelvre való kiterjesztést. A szabályalapú gépi fordítás módszer előnye, hogy meglévő elemző rendszerek mellett magas pontosságot eredményez, azonban ha a szótár kicsi, a fedése általában alacsony, vagyis kevés dolgot tud lefordítani.

Nagyon leegyszerűsítve, a szabályalapú gépi fordítás működése hasonlít ahhoz, amikor egy felnőtt ember nyelvet tanul. Külön megtanuljuk a szavakat egy szótár segítségével és

külön megtanuljuk a nyelvtani szabályokat egy nyelvtankönyvből, majd a megszerzett tudás segítségével próbálunk lefordítani egy szöveget. Ha kevés a szókincsünk, akkor kevés dolgot tudunk lefordítani, ha hiányos a nyelvtani tudásunk, akkor sok hibát fogunk vétetni.

Magyarországon csupán egyetlen neves képviselője a szabályalapú gépi fordító módszernek, a webforditas.hu¹ online fordító rendszer.



1. ábra. Vauquois-háromszög

Példaalapú gépi fordítás

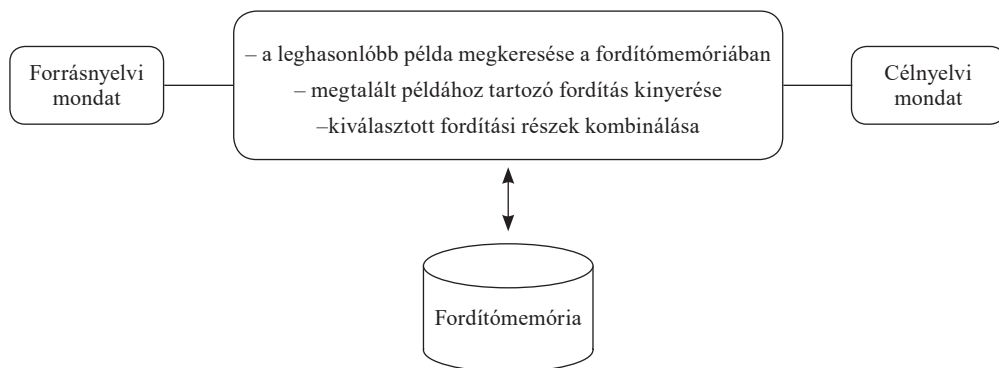
A példaalapú gépi fordítás (*Example Based Machine Translation*) (Somers 1999) alapja az, hogy a már emberek által lefordított szövegeket felhasználjuk. Az előre lefordított szövegszegmenseket mint példákat eltároljuk egy fordítómemóriában. Így a fordítási memóriába általában a gyakran előforduló lefordított kifejezéseket és mondatokat mentik el. A rendszer ezekből választja ki a fordítandó szöveghez a hozzá leghasonlóbb kifejezéseket vagy mondatokat, majd a kiválasztott részek egyesítésével állítja elő a fordítást (lásd 2. ábra). A rendszer előnye, hogy azokat a szövegrészeket, amelyek szerepelnek a fordítómemóriában, magas pontossággal le tudja fordítani. Azonban ha egy szövegrészlet nincsen benne a fordítómemóriában, a rendszer pontatlanul vagy egyáltalán nem kezeli.

A példaalapú gépi fordítás középpontjában a fordítómemória áll, amit külön is lehet kezelni. A fordítómemória önmagában alkalmazható más módszereknél is.

A példaalapú gépi fordító módszernek egy híres hazai példája a memoQ².

¹ <http://www.webforditas.hu/>

² <https://www.memoq.com/>



2. ábra. Fordítómemória működési elve

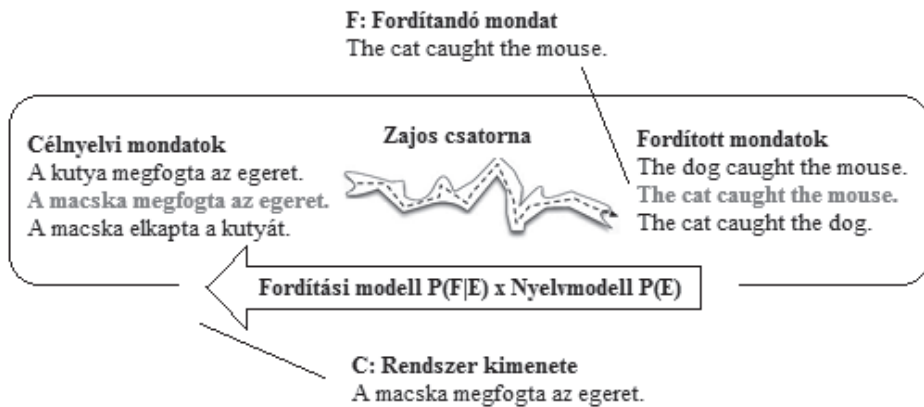
A statisztikai gépi fordítás

A statisztikai gépi fordítás (*Statistical Machine Translation*) (Koehn 2010) párhuzamos szövegkorpuszon alapszik. A példaalapú gépi fordítás kialakulásával egyidejűleg alakult ki a statisztikai gépi fordítás is. Hasonlóan a példaalapú módszerhez, ez is az emberek által előre lefordított szövegek felhasználásán alapszik. A párhuzamos korpuszok előre lefordított nagy mennyiségű szövegek. A rendszer a párhuzamos korpuszt elemezve felállít egy szótárt, illetve statisztikai módszereken alapuló megfigyeléseket tesz, amelyek alapján fordít. A probléma megoldásához a beszédtechnológiában használt Shannon-féle zajoscsatorna-modellt vették alapul (Shannon, 1948). A rendszer modelljét (lásd 3. ábra) az alábbi módon írhatjuk le: adott egy F mint forrásnyelv, amit szeretnénk lefordítani E célnyelvre. Amit biztosan ismerünk, az a lefordítandó szövegünk. A fordítás úgy történik, hogy vesszük az összes lehetséges célnyelvi mondatot, majd mintha átéreznénk egy zajos csatornán, majd az így kapott kimenetet összehasonlítanánk a fordítandó mondatunkkal, és amelyek 'zajos mondat' a legjobban hasonlít rá, vagyis amelyek a legnagyobb valószínűséggel a fordítás lesz (Bayes-féle becsléssel számítjuk ki a P valószínűséget), az lesz a rendszer kimenete (C).

A statisztikai gépi fordító rendszer felépítéséhez szükségünk van egy fordítási modellre, egy nyelvmodellre, valamint egy dekódoló rendszerre, amely összeköti a fordítási modellt a nyelvmodellel és megtalálja a legvalószínűbb fordítást a forrásnyelvi szegmens alapján. A fordítási modell felel a fordítás tartalomhűségéért (*adequacy*), míg a nyelvmodell felel a célnyelvi mondat folyékonyságáért vagy nyelvhelyességéért (*fluency*). A fordítási szegmensek, amelyek egyben meghatározzák a statisztikai gépi fordító rendszer fajtáját is, lehetnek szavak (szóalapú statisztikai gépi fordító rendszer), kifejezések (kifejezésalapú statisztikai gépi fordító rendszer), tulajdonsághalmazok (faktoralapú statisztikai gépi fordító rendszer) és generatív szabályok (szintaxisalapú statisztikai gépi fordító rendszer). A módszer legnagyobb előnye, hogy nem kell ismerni az adott nyelveket, amelyekre alkalmazni szeretnénk a rendszert, bármilyen két nyelvet be lehet tanítani. Továbbá nincsen szükség emberi közreműködésre, valamint az internet széleskörű elterjedésének köszönhetően igen nagyszámú többnyelvű digitális szöveg áll a rendelkezésünkre. Hátránya, hogy ha kevés a tanítóanyag, akkor nem lesz jó a modell minősége, illetve ha a tanítóanyag minősége rossz, akkor a betanított modell sem lesz jó, vagyis nagyban függ a modell a korpusz minőségétől. Továbbá fennáll az az eset is, hogy téma (*domain*) specifikusan tanítjuk be a modellt, ez esetben a gépi fordító rendszer csak az adott témán belül képes megfelelő minőségben fordítani.

Ha nagyon leegyszerűsítjük, a statisztikai gépi fordítás működése hasonlít ahhoz, amikor egy kisgyermek tanulja a nyelvet. Ha sokszor megmutatjuk neki, hogy mi az asztal, egy idő után megtanulja az asztal fogalmát. Ha kevés dolgot mutatunk neki, kevés dolgot fog megtanulni, vagy ha félrevezetjük, és olyan dolgokra is azt mondjuk, hogy asztal, ami nem az, akkor rosszul fogja megtanulni az asztal fogalmát. Például a Google fordító egyik látványos hibája volt, hogy a 'The dog was bitten by the postman.' mondatot úgy fordította, hogy 'A kutya megharapta a postást.' Ez jól jellemzi a statisztikai módszer hibáját, miszerint a 'kutya', 'harap' és 'postás' összetételben az a mondat a leggyakoribb, hogy 'A kutya megharapta a postást.'

Az egyik legközismertebb statisztikai gépi fordító rendszer a már említett Google fordító³ volt 2016 előtt, ma a Microsoft Bing fordítója⁴ még statisztikai módszeren alapszik.



3. ábra. Zajoscatorna-model

A hibrid gépi fordítás

A tényleges alkalmazások során, a jobb eredmény érdekében gyakran ötvözik a különböző gépi fordító módszerek előnyeit. A hibrid gépi fordítás (*Hybrid Machine Translation*) célja, hogy egy adott típusú gépi fordító módszer minőségét és pontosságát javítsa más gépi fordító módszerek integrálásával. Ilyen például, ha a szabályalapú döntéseknél a rendszer figyelembe vesz statisztikai eredményeket vagy a statisztikai fordításnál szabályrendszert is alkalmaznak. Vagy akár a pontosabb fordítás érdekében, mind a szabály, mind a statisztikai módszernél alkalmaznak fordítómemóriát.

A neurális gépi fordítás

A neurális gépi fordítás (*Neural Machine Translation*) a mesterséges neurális hálózatokon alapszik. A mesterséges neurális hálózat egy gépi tanulási módszer, amely adott számú bemenet alapján végez predikciót. A rendszer hátránya, hogy a tanításhoz nagy mennyiségű tanítóanyagra és nagy kapacitású számítógépre van szükség, amelynek beszerzése nagy költséggel jár és komoly szaktudást igényel. Az előnye hasonló a statisztikai gépi fordító rend-

³ <https://translate.google.hu/>

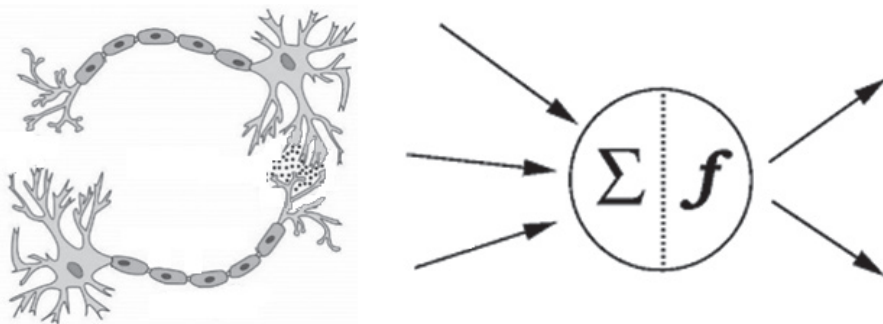
⁴ <https://www.bing.com/translator>

szeréhez, miszerint nem kell ismerni az adott nyelveket a modell tanításához. A legnagyobb előnye azonban, hogy képes tanulni a saját hibáiból és a tanítás során képes fejlődni.

A számítógépek, de főleg a grafikus kártyák teljesítményének növekedésével gyors ütemben fejlődött a neurális hálózat módszere is. Az utóbbi években, számtalan tudományágban, többek között a gépi fordítás területén is érték el áttörő eredményeket a neurális hálózat előnyeinek kiaknázásával. A nagyobb cégek és kutatások már átállították a fókuszot a neurális módszerekre. Többek között a Google (Wu és társai 2016) és a Microsoft (Microsoft 2018) is cserélik le a gépi fordító rendszerüket neurális alapokra. A következő fejezetben bővebben fejtem ki a neurális gépi fordítás működését.

A neurális gépi fordítás

A neurális hálózat elméletét már 1943-ban megfogalmazták (McCulloch és Pitts 1943). A neurális gépi fordítás alapja a mesterséges neurális hálózat, amely mesterséges neuronokból épül fel. A mesterséges neuronok az idegsejteket modellálják. Az emberi agy több milliárd idegsejtet tartalmaz, ezek működésének hatékonysága ihlette a mesterséges neuronok létrehozását. A neuron egy információfeldolgozó egység, amely bemeneti adatokból egyszerű számításkkal generál kimeneti adatot (lásd 4. ábra). A számítás első része lehet a bemeneti értékek egyszerű súlyozott összegzése (Σ), de lehet komplexebb művelet is. A számítás második része, ami esetenként elmaradhat, egy átviteli függvény (f), ami a súlyozott bemenetekből állítja elő a kimeneti értékeket. Az átviteli függvény általában egy küszöb vagy egy szigmoid-függvény, de ez is lehet komplexebb (pl.: Gauss-féle függvény).



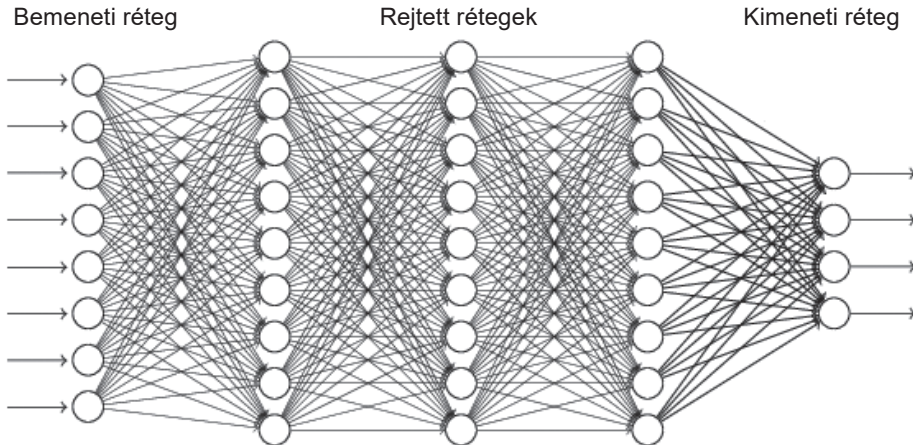
4. ábra. Neuron modellezése

A neurális hálózat

A mesterséges neuronok összekapcsolásával jön létre a mesterséges neurális hálózat. A neurális hálózat három funkcionálisan és strukturálisan elkülöníthető rétegre osztható (lásd 5. ábra):

1. *Bemeneti réteg*: a hálózat a bemeneti rétegen keresztül kapja az információt a külvilágtól, amit módosítás nélkül továbbít a hálózat további részébe.
2. *Kimeneti réteg*: a kimeneti réteg állítja elő a hálózat eredményét. Feladattól függően változik a kimeneti réteg. Például osztályozási feladat esetében a kimeneti rétegben annyi neuron van, ahány osztály.
3. *Rejtett rétegek*: a rejtett rétegek a bemeneti és a kimeneti rétegek között helyezkednek el. A legegyszerűbb neurális hálózat egy rejtett réteggel rendelkezik. Minél több rejtett réteg van a hálózatnak, annál jobban növekszik az absztrakciós képessége és annál

összetettebb feladatokat képes megoldani. A rejtett rétegek számának növelésével „mélyül” a hálózat, ezért a ’mély neuronhálózat’ kifejezés a rétegek számára utal, a ’mély tanulás’ (*deep learning*) kifejezés pedig a több rejtett réteggel rendelkező neurális hálózat tanítását jelenti.



5. ábra. Neurális hálózat (Nielsen 2015)

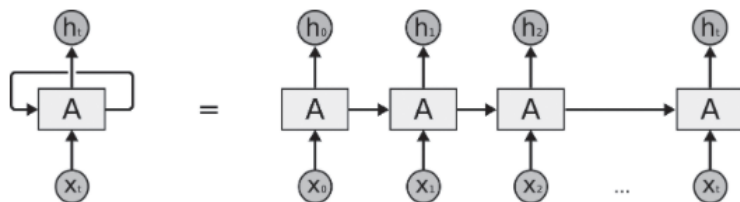
A neurális hálózat topológiája alapján lehet *előre*csatolt és *vissza*csatolt. Az 5. ábra egy előre-csatolt hálózatot ábrázol. A visszacsatolt hálózat esetében nem csak előre csatolások, hanem a rétegekben belül és a rétegek között visszacsatolások is lehetnek.

A neurális hálózat módszer egyik legnagyobb előnye, hogy képes tanulni a saját hibájából. A hálózat a rendszer hibáját egy veszteségfüggvénnyel és várt kimenetek segítségével számolja ki. A veszteségfüggvény lehet átlagos négyzetes eltérés (*mean squared error*), különböző kereszt-entrópia (*cross-entropy*) függvények stb. A hiba meghatározásánál először a kimeneti hibákat számoljuk ki, majd a hiba-visszaterjesztés alkalmazásával kiszámoljuk a rejtett rétegek hibáit is. Amikor a rendszer a hibákat kiszámolta, az egyes rejtett rétegekben a súlyokat úgy állítja át, hogy az adott rétegre számolt hibamérték csökkenjen. Ezt a hiba-visszaterjesztési folyamatot a tanítás során többször elvégezi, optimális esetben addig, amíg a hibamérték minimális nem lesz.

Ha leegyszerűsítjük, a neurális hálózat működése hasonlít ahhoz, amikor egy ember zongorázni tanul. Az elején még sok hibával játssza el a darabot, de a folyamatos gyakorlás (hibajavítás) után az akkordok szépen összeállnak egy szép zeneművé.

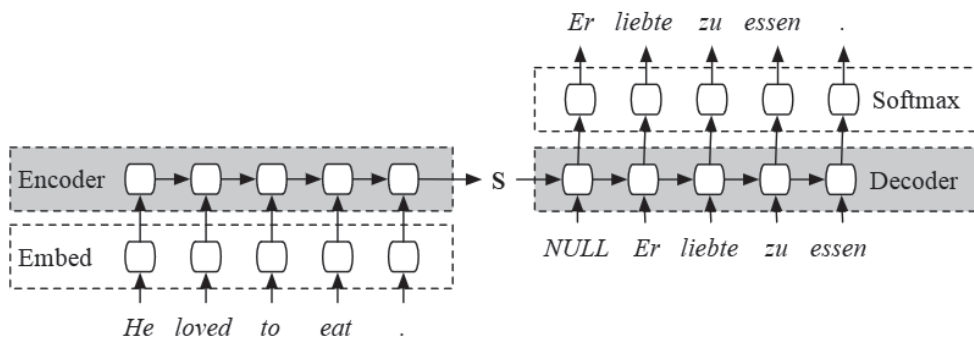
A neurális hálózat a gépi fordításban

A neurális gépi fordítás rekurrens neurális hálózatot (*Recurrent Neural Network – RNN*) használ. Ahogy a 6. ábra mutatja, az RNN egy visszacsatolt hálózat, amely a bemeneti szöveg egységein iterálva számol ki egy súlyozott értéket a bemeneti szövegre. Az ábrán az x a bemenet, a h a kimenet és az A egy rejtett rétegben lévő neuron. A szöveg egysége lehet szó, szórészeslet vagy akár betű (karakter) is.



6. ábra. A rekurrens neurális hálózat működése (Olah 2015)

A neurális gépi fordítás (Cho és társai 2014b) kettő RNN-hálózatból épül fel (lásd 7. ábra szürkével jelölt részei): egy encoder és egy decoder. Az encoder felel azért, hogy a változó hosszúságú bemeneti szövegből leképezzen egy állandó hosszúságú vektor reprezentációt, amivel majd az RNN dolgozik. A decoder feladata az, hogy az RNN által használt fix hosszúságú vektor reprezentációból kimeneti szöveget generáljon.



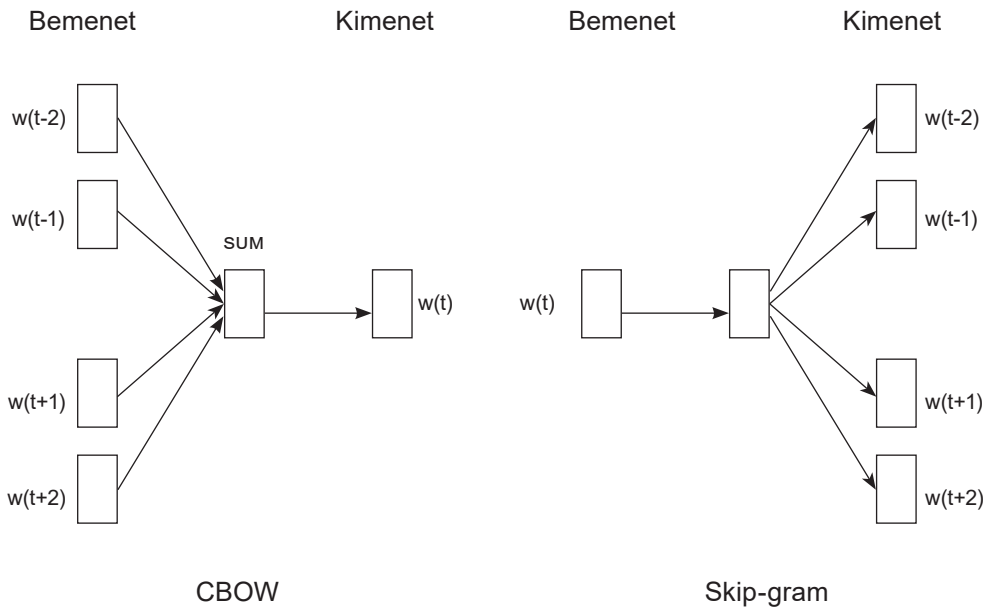
7. ábra. A neurális gépi fordítás működése (Merity 2016)

A 7. ábra alapján láthatunk egy-egy plusz réteget mind a bemenetnél (Embed), mind a kimenetnél (Softmax). Ezek a rétegek az úgynevezett szóreprezentáció rétegek. Az RNN számokkal dolgozik. A szóreprezentáció lényege, hogy az RNN nem közvetlenül a nyers bemeneti szöveget dolgozza fel, hanem a szöveg egységeiből (jelen esetben szavak) szövektorokat készít és ezt adja oda az RNN-nek. A szövektor előállítását kezdetben az úgynevezett one-hot vektor módszerrel történt. A one-hot vektor egy olyan vektor, amiben csak 0-ák állnak, kivéve egyetlen elemet, ami egy 1-es. A vektor hossza egyenlő a neurális gépi fordító rendszer tanításához használt tanítóanyagból létrehozott szótár hosszával. A vektorban szereplő 1-es indexe (helyének sorszáma) megegyezik az általa reprezentált szó szótárbeli indexével (helyének sorszáma). Azonban ez a módszer mindössze egy technikai módszer a szavak reprezentációjára, információtartalma nincsen. Ezért később a one-hot vektor módszerét felváltotta a szóbeágyazás (word embedding) módszere, ami már szemantikai információt is tartalmaz.

A szóbeágyazás (Mikolov–Sutskever–Chen–Corrado–Dean 2013; Mikolov–Chen–Corrado–Dean 2013) azon az elméleten alapszik, hogy a hasonló jelentésű szavakat hasonló környezetben használjuk. A szóbeágyazás módszerében a lexikai elemek egy valós vektortérben egy-egy pontnak felelnek meg, amelyek konzisztensen helyezkednek el az adott térben. Ebben

az adott térben a szemantikailag közel álló szavak közel esnek egymáshoz, míg a jelentésben távol álló szavak távol vannak egymástól. A szemantikai hasonlóságot a két pont közötti távolsággal írhatjuk le. A modell tanításához újabb neurális hálózatra van szükségünk. Kétféleképpen taníthatjuk a modellt. Első eset, amikor a neurális hálózat bemenete egy szó, és a kimenete a bementi szó fix méretű környezetének szavai, ezt a Skip-gram modell (lásd 8. ábra) segítségével tudjuk betanítani. A másik lehetőség, hogy a hálózat bemenete egy szó fix környezetének szavai, és a kimenet az adott szó, aminek a környezetét vizsgáljuk. Ez utóbbi modellt a CBOW (*continuous bag-of-words*) modell (lásd 8. ábra) segítségével lehet betanítani.

Az így létrejött modellek segítségével a rendszer minden szóhoz hozzárendeli a hozzá tartozó szóreprezentációját, ami egy szövektor. Ez a szövektor lesz a neurális gépi fordítás bemenete, ami egyben már szemantikai információt is tartalmaz.



8. ábra. CBOW és Skip-gram működése (Mikolov–Chen–Corrado–Dean 2013)

A neurális gépi fordítás további erőssége, hogy ugyanazt a modellt több nyelvre is be lehet tanítani. A gyakorlatban úgy működik ez, hogy elkezdik két nyelvre betanítani egy modellt, és egy idő után megváltoztatják a forrásnyelvet. Például magyar–kínai mondatpárból kevés van, ezáltal kevés tanító anyag is van, angol–kínai mondatpárból viszont sokkal több van. Ezért angol–kínai mondatpárokkal kezdik tanítani a modellt, majd egy idő után magyar–kínai mondatokkal. Ennek az következménye, hogy mind az angol, mind a magyar forrásnyelv ugyanabba a térbe képződik le, ezzel létrejön a magyar és az angol szavak között a kapcsolat. Így, bár kevés magyar–kínai tanítóanyag van, mégis jó minőségben tud fordítani a gép magyar nyelvről kínai nyelvre, mert az angol szavak ebben a modellben kiegészítik a magyar szavakat. Ezzel a módszerrel gyakorlatilag bármilyen nyelvről bármilyen nyelvre tud fordítani a gép, csak elég tanítóanyagot kell biztosítani neki. Ha az egyik forrásnyelvből kifogy a tanítóanyag, elkezdjük másik forrásnyelv segítségével továbbtanítani. A Google fordító ennek a módszernek a segítségével hidalja át a kisebb nyelvek közötti fordítás nehézségét.

Véleményem szerint jelenleg a Google fordító a legelőrehaladottabb a neurális gépi fordítás terén. A birtokában lévő óriási mennyiségű adat és anyagi háttér biztosítja a folyamatos fejlődését, aminek következménye, hogy emberi szemmel látványosan jobb minőségű fordításokat produkál, mind az egyszerű, mind az összetett mondatok esetében és több nyelven, többek között magyar nyelven is.

Összegzés

A gépi fordítás fejlődése nagy utat tett meg. A története a szabályalapú módszerekkel kezdődött, majd ahogy egyre több fordítás született, azok felhasználásával alakultak ki a példaalapú és a statisztikai módszerek. A neurális hálózat módszerét már 1943-ban kidolgozták, azonban a hálózat előnyeit csak a számítógépek fejlődésével tudták kiaknázni. Az igazi áttörés akkor történt, amikor a grafikus kártyák fejlődése elérte azt a teljesítményt, hogy képesek lettek nagy számításigényű feladatok elvégzésére. Számptalan tudományágban értek el áttörő eredményt a neurális hálózat segítségével, és ez a trend a gépi fordítás szakterületét is elérte. A neurális hálózat egyik hátránya és egyben előnye is, hogy rendkívüli nagy mennyiségű adat szükséges a tanításához. Ha nincs elég adat, akkor gyenge a minősége, azonban megfelelő mennyiségű tanítóanyaggal rendkívüli látványos eredményeket lehet vele elérni. A neurális hálózat az absztrakt rétegeiben képes mintákat felismerni és megtanulni. A legnagyobb előnye mégis az, hogy képes tanulni a saját hibáiból, és ha elég nagy mennyiségű szöveg áll rendelkezésünkre, akkor rendkívüli finomsággal képes megtanulni mintákat. Ezzel több szabályosságot tud felismerni, mint amit egy ember képes megfogalmazni a szabályalapú rendszereknél, és míg a statisztikai módszer a szabályokat csak a szavak adott hosszúságú kontextusából tudja előállítani, addig a neurális módszer képes a szövegben sokkal nagyobb távolságban lévő összefüggéseket felismerni. Ezen tulajdonságai miatt válik egyre pontosabbá és népszerűbbé a neurális gépi fordítás. Jelentősége oly nagymértékű, hogy a nagyobb cégek, mint például a Google is leváltotta a statisztikai gépi fordító rendszerét neurális gépi fordításra.

Yang Zijian Győző

IRODALOM

- Cho, K. – Merriënboer, B. v. – Bahdanau, D. – Bengio, Y (2014a): On the Properties of Neural Machine Translation: Encoder–Decoder. *Proceedings of SSSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*, Association for Computational Linguistics, Doha, Qatar, 103–111.
- Cho, K. – Merriënboer, B. v. – Gulcehre, C. – Bahdanau, D. – Bougares, F. – Schwenk, H. – Bengio, Y (2014b): Learning Phrase Representations using RNN Encoder–Decoder. *Proceedings of the 2014 Conference on Empirical Methods on Natural Language Processing*, Association for Computational Linguistics, Doha, Qatar, 1724–1734.
- Hutchins, W. J. – Somers, H. L (1992): *An Introduction to Machine Translation*. London: Academic Press.
- Koehn, P. (2010): *Statistical Machine Translation*. Cambridge: Cambridge University Press.
- McCulloch, W. S. – Pitts, W. (1943): A Logical Calculus of the ideas immanent in Nervous Activity. *The bulletin of mathematical biophysics*, vol. 5, Chicago: The University of Chicago Press, 115–133.
- Merity, S. (2016. november 17.): *Peeking into the neural network architecture used for Google's Neural Machine Translation*. [Letöltés dátuma: 2018. 07 10], Forrás: Peeking into the neural network

- architecture used for Google's Neural Machine Translation: https://smerity.com/articles/2016/google_nmt_arch.html
- Microsoft. (2018): *Machine Translation*. [Letöltés dátuma: 2018. 07 10], forrás: Microsoft: <https://www.microsoft.com/en-us/translator/mt.aspx#nnt>
- Mikolov, T. – Chen, K. – Corrado, G. – Dean, J. (2013): Efficient estimation of word representations in vector space. *International Conference on Learning Representations*, Scottsdale, USA.
- Mikolov, T. – Sutskever, I. – Chen, K. – Corrado, G. – Dean, J. (2013): Distributed Representations of Words and Phrases and their Compositionality. *Proceedings of the 27th Annual Conference on Neural Information Processing Systems*, vol 2, Cambridge: MIT Press Cambridge, 3111–3119.
- Nielsen, M. A. (2015): *Neural Networks and Deep Learning*. Determination Press, <http://neuralnetworksanddeeplearning.com/>, [Letöltés dátuma: 2018. 07 22]
- Olah, C. (2015. augusztus 27.): *Understanding LSTM Networks*. [Letöltés dátuma: 2018. 07 10], Forrás: Understanding LSTM Networks: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Shannon, C. E. (1948): A Mathematical Theory of Communication. *Bell System Technical Journal*, 27/3, 379–423.
- Somers, H. (1999): Review Article: Example-based Machine Translation. *Machine Translation*, 14/2, 113-157.
- Wu, Y. – Schuster, M. – Chen, Z. – Le, Q. V. – Norouzi, M. – Macherey, W. – Krikun, M. – Cao, Y. – Gao, Q. – Macherey, K. – Klingner, J. – Shah, A. – Johnson, M. – Liu, X. – Kaiser, Ł. – Gouws, S. – Kato, Y. – Kudo, T. – Kazawa, H. – Stevens, K. – Kurian, G. – Patil, N. – Wang, W. – Young, C. – Smith, J. – Riesa, J. – Rudnick, A. – Vinyals, O. – Corrado, G. – Hughes, M. – Dean, J. (2016): Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. *Technical Report*, arXiv preprint arXiv:1609.08144, <http://arxiv.org/abs/1609.08144>, . [Letöltés dátuma: 2018. 07 22].