

امروزه با گسترش فزاینده شبکه های جهانی و اسناد اینترنتی، نیاز به ترجمه این اسناد از یک زبان به زبان دیگر روز به روز در حال افزایش است. در حال حاضر نیاز به ترجمه تا حدودی رفع شده و تمرکز حوزه ترجمه بیشتر بر روی افزایش کیفیت ترجمه می باشد. این افزایش کیفیت در جفت زبان هایی مانند زبان های انگلیسی و فرانسوی که از نظر نحوی و ساختواژی شبیه هم هستند، پیشرفت خوبی داشته است. ولی در مقابل، زبان هایی مانند فارسی و عربی که از نظر نحوی و ساختواژی و همچنین ترتیب کلمات با زبان هایی نظیر انگلیسی تفاوت های زیادی دارند، کار دشوار و پیچیده تری می باشد.

به منظور افزایش کیفیت ترجمه ماشینی انگلیسی به فارسی و به طور خاص مترجم ماشینی مبتنی بر قاعده، مترجم ماشینی با معماری ترکیبی طراحی شده است که با استفاده از مترجم ماشینی آماری، نتایج مترجم ماشینی مبتنی بر قاعده را بهبود می دهد. این معماری بر پایه مترجم ماشینی مبتنی بر قاعده طراحی و ساخته شده است و به منظور بهبود نتایج، خروجی مترجم ماشینی قاعده مند و آماری را ترکیب کرده، و در این راستا از یک رمزگشای یکنواخت با پیچیدگی زمانی چندجمله ای استفاده می کند. مترجم های ماشینی قاعده مند عمل ترجمه را بر اساس مجموعه ای از قواعد زبانی انجام می دهند که معمولاً نتایج آنها از نظر جابجایی کلمات و ساختار نحوی، کیفیت بهتری نسبت به نتایج مترجم های آماری دارند. ولی عملکرد این مترجم ها در زمینه انتخاب لغات مناسب و روانی ترجمه، ضعیف تر از مترجم های ماشینی آماری است. در واقع مترجم های ماشینی قاعده مند دانش محور بوده و عمل ترجمه را بر اساس دانش زبان شناسی انجام می دهند و به همین دلیل ترجمه آنها از نظر ساختار و رعایت نکات نحوی و ساختواژی بهتر از مترجم های آماری می باشد. در مقابل مترجم های آماری داده محور بوده و عمل ترجمه خود را بر اساس دادگان از پیش آماده شده انجام می دهند، از این رو ترجمه آنها به ترجمه انسانی نزدیک تر است و کلمات و عبارات روان تر و سلیس تر ترجمه می شوند.

از این رو در معماری ارائه شده، ترجمه اولیه بوسیله مترجم ماشینی قاعده مند صورت می گیرد تا نکات نحوی و ساختواژی در آن لحاظ شود، سپس با استفاده از مترجم ماشینی آماری و ترجمه مناسب تر و روان تر کلمات و عبارات، ترجمه آن بهبود داده می شود. به منظور این بهبود، ترتیب کلمات در ترجمه نهایی بر اساس ترجمه مترجم ماشینی قاعده مند صورت می گیرد. سپس عمل ترجمه و انتخاب لغات توسط یک رمزگشای یکنواخت، با در نظر گرفتن ترجمه های کاندیدای ارائه شده توسط مترجم قاعده مند و آماری و همچنین با استفاده از مدل زبانی، انجام می شود. آزمایش های انجام شده نشان می دهند که کیفیت نتایج بدست آمده از معماری ترکیبی بر روی دادگان تست در معیار بلو، تقریباً ۵ واحد بهتر از نتایج مترجم ماشینی قاعده مند است. همچنین کیفیت این نتایج نسبت به نتایج مترجم ماشینی آماری در معیار بلو، یک واحد بهتر است.