

فهرست مطالب

1- مقدمه	1
1-1- طرح پیشنهادی	2
2-1- مقالات مستخرج از پایان نامه	4
3-1- ساختار پایان نامه	5
2- مرور ادبیات	7
1-2- تعاریف پایه زبان شناسی	7
1-1-2- ریشه یابی	7
2-1-2- برچسب زنی بخشهای سخن (POS)	7
3-1-2- درخت تجزیه	8
4-1-2- پیکره تک زبانه	9
5-1-2- پیکره موازی	10
2-2- ترجمه ماشینی	11
3-2- مشکلات موجود در ترجمه ماشینی	13
1-3-2- ابهام	13
2-3-2- تفاوت‌های ساختاری	14
3-3-2- تفاوت‌های لغوی	14
4-2- روش‌های موجود برای ترجمه ماشینی	14
1-4-2- رویکرد تبدیل	15
2-4-2- Interlingua	16
3-4-2- رویکرد مستقیم	17
4-4-2- رویکردهای برپایه پیکره	19
5-2- ترجمه ماشینی آماری	21
1-5-2- مدل زبانی با استفاده از N-gram ها	24
2-5-2- مدل ترجمه	27
3-5-2- روش‌های مختلف ترجمه ماشینی آماری	30
6-2- ارزیابی سیستم‌های ترجمه	33
1-6-2- سیستم ارزیابی BLEU	34
2-6-2- سیستم ارزیابی NIST	35
7-2- پیشینه تحقیق	35
1-7-2- استفاده از اطلاعات زبان‌شناسی جهت بهبود ترجمه ماشینی آماری	36

47	2-7-2- ترجمه ماشینی آماری انگلیسی به فارسی
53	8-2- چالش‌های موجود
53	1-8-2- نبود پیکره دوزبانه مناسب
53	2-8-2- محدودیت در مورد کار با پیکره‌های بزرگ
53	3-8-2- تفاوت ساختاری بسیار زیاد بین زبان فارسی و انگلیسی
54	9-2- خلاصه فصل
57	3- روش پیشنهادی
57	1-3- بازآرایی نحوی جملات
58	1-1-3- قوانین بازآرایی
63	2-3- مدل بر مبنای فاکتور
64	1-2-3- ریشه‌یاب فارسی
69	2-2-3- برچسب زنی ادات سخن
69	3-3- خلاصه فصل
71	4- پیاده سازی و ارزیابی
71	1-4- مقدمه
72	2-4- تولید پیکره آموزشی FEP
74	3-4- حذف حروف و علائم زائد
75	4-4- ایجاد سیستم مترجم آماری
76	5-4- مترجم آماری انگلیسی به فارسی بر مبنای بازآرایی نحوی
79	6-4- ایجاد ترجمه ماشینی آماری بر مبنای فاکتور
88	7-4- خلاصه فصل
90	5- نتیجه‌گیری و کارهای آتی
90	1-5- نتیجه‌گیری
90	2-5- کارهای آتی
93	6- مراجع