

## حاشیه نویسی معنایی خودکار اسناد وب توسط استخراج قوانین پیوند در متن

بهنام حاجیان<sup>۳۵</sup>، دکتر کامران زمانی فر<sup>۳۶</sup>

### چکیده

پس از ظهور وب معنایی و قابل درک و استنتاج برای کامپیوتر توسط آقای برنرزی، صفحات وب موجود میبایست توسط استانداردهای مطرح شده توسط ائتلاف جهانی وب با آنتولوژی مربوطه حاشیه نویسی معنایی گردند تا دامنه معنایی کلمات متن مشخص گردد و مشکلات مطرح شده در زمینه بازیابی و جستجوی اطلاعات در محیط وب حل شود. ولی در این میان به مشکل بزرگ و عمده ای در محیط وب معنایی بر خواهیم خورد و آن انجام عمل حاشیه نویسی معنایی برای صفحات وب متداول است. در حال حاضر این عمل توسط عامل انسانی بصورت دستی انجام می گیرد. با توجه به کمبود دقت انسان در انجام این عمل پیچیده، وقتگیر و پرهزینه و توجه وسیع به هوش ماشینی و الگوریتم های داده کاوی در سال های اخیر تصمیم به ارایه الگوریتمی جهت خود کار نمودن حاشیه نویسی معنایی وب توسط ماشین نموده ایم. بدین جهت توجه ما به وابستگی معنایی بین کلمات معطوف شده است. اجرای این محیط بوسیله ابداع الگوریتمی بر مبنای الگوریتمهای آموزش ماشین جهت استخراج قوانین پیوند بین کلمات موجود در متونی که قبلا توسط انسان حاشیه نویسی شده انجام گرفته است.

الگوریتم فوق شامل دو مرحله است: - مرحله آموزش بر مبنای صفحات حاشیه نویسی شده توسط انسان جهت استخراج قوانین بصورت خود کار توسط ماشین - مرحله بکارگیری قوانین استخراج شده جهت حاشیه نویسی معنایی خودکار متن و تعیین آنتولوژی کلمات موجود در متن و رفع ابهام معانی آنها میباشد.

### کلمات کلیدی

آنتولوژی، حاشیه نویسی معنایی، وب معنایی، قوانین پیوند، متن کاوی، آموزش ماشین

<sup>35</sup> دانشجوی کارشناسی ارشد دانشگاه اصفهان گروه مهندسی کامپیوتر be\_hajian@mehri.ui.ac.ir

<sup>36</sup> عضو هیئت علمی گروه کامپیوتر دانشگاه اصفهان zamanifar@eng.ui.ac.ir