

<p>کد مسئله : SI-20</p>	<p>فرم درخواست تعریف پروژه (RFP)</p>	
<p>عنوان مسئله : توسعه سرور هوش مصنوعی</p>		
<p>تعریف مسئله، ضرورت انجام و اهداف طرح:  سرور هوش مصنوعی یا AI سرور، سیستمی است که به طور ویژه برای اجرای اپلیکیشن‌های هوش مصنوعی طراحی شده است. این سرورها دارای قدرت پردازشی بالایی هستند و به یک یا چند پردازنده‌های گرافیکی (GPU) قوی، شتاب‌دهنده‌هایی اختصاصی هوش مصنوعی مانند واحدهای پردازش tensor یا کارت‌های گرافیک مجهز هستند.</p>		
<p>مشخصات فنی و عملیاتی:</p>		
<p>خروجی‌های مورد انتظار (دستاوردهای فنی و تولیدات علمی):</p> <ul style="list-style-type: none"> <li>● زیرساخت سخت‌افزاری <ul style="list-style-type: none"> <li>○ ساخت شتاب‌دهنده سخت‌افزاری با کاربرد پردازش گرافیکی در سرور AI</li> <li>○ طراحی ماژول RAM</li> <li>○ دستیابی به معماری بهینه برای کلاسترینگ سرورهای AI</li> </ul> </li> <li>● زیرساخت نرم‌افزاری <ul style="list-style-type: none"> <li>○ توسعه و سفارشی‌سازی سیستم‌عامل سرور AI</li> <li>○ دستیابی به معماری میان‌افزار BMC با هدف امن‌سازی سرورهای AI</li> <li>○ کلاسترینگ Multi Node GPU مبتنی بر کانتینر</li> <li>○ مانیتورینگ کلاستر به منظور ارائه عملکرد زیرساخت</li> <li>○ تست و بنچمارک به منظور دستیابی به حداکثر کارایی زیرساخت</li> </ul> </li> <li>● پلتفرم <ul style="list-style-type: none"> <li>○ طراحی و سفارشی‌سازی پلتفرم جهت توسعه سرویس‌های AI</li> <li>○ دستیابی به مدل پلتفرم توسعه مستمر بر مبنای ابزارهای CI/CD</li> <li>○ دستیابی به پلتفرم مدیریت داده</li> <li>○ دستیابی به پلتفرم مدیریت مدل</li> </ul> </li> </ul> <p>توسعه مکانیزم اتصال به خدمات ابری</p>		
<p>محدودیت‌ها و قیود:</p>		